

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:43:59 ; Search time 11 Seconds  
(without alignments)  
1315:931 Million cell updates/sec

Title: US-10-027-059a-1

Perfect score: 1889

Sequence: 1 MADADSGFLAHPLEPAK.....CDHPWSICFLSLFLSLGNG 349

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1767	93.5	518	1	TXB5_HUMAN
2	1712	90.6	518	1	TXB5_MOUSE
3	1130.5	59.8	545	1	TXB4_HUMAN
4	865.5	45.8	173	1	TXB4_MOUSE
5	658.5	34.9	423	1	TXB4_CABEL
6	653.5	34.6	414	1	TXB3_CHICK
7	652.5	34.5	361	1	TXB1_CHICK
8	649.5	34.4	701	1	TXB2_MOUSE
9	635.5	33.6	436	1	TXB5_HUMAN
10	633.5	33.5	702	1	TXB6_HUMAN
11	629.5	33.3	540	1	TXB6_MOUSE
12	626	33.1	398	1	TXB1_HUMAN
13	626	33.1	988	1	OMB_DROME
14	624.5	33.1	742	1	TXB3_HUMAN
15	616.5	32.6	251	1	TX20_HUMAN
16	616.5	32.6	297	1	TX20_MOUSE
17	606	32.1	455	1	VEGT_XENLA
18	603.5	31.9	613	1	TX18_MOUSE
19	601	31.8	501	1	TX18_HUMAN
20	591.5	31.3	473	1	TX16_BRARE
21	589	31.2	602	1	TX15_MOUSE
22	587.5	31.1	184	1	TXB1_MOUSE
23	583.5	30.9	181	1	TXB3_MOUSE
24	575	30.4	681	1	BRC2_HALRO
25	558	29.5	177	1	TX13_MOUSE
26	549	29.1	346	1	TX12_CABEL
27	544.5	28.8	436	1	BRAC_MOUSE
28	542.5	28.7	435	1	BRAC_HUMAN
29	541.5	28.7	432	1	BRAC_XENLA
30	541.5	28.7	544	1	H15_DROME
31	536.5	28.4	423	1	BRAC_BRARE
32	536.5	28.4	681	1	TBRL_MOUSE
33	536.5	28.4	682	1	TBRL_HUMAN

RESULT 1  
TXB5\_HUMAN  
ID TXB5\_HUMAN STANDARD: PRT: 518 AA.  
AC O9593; G9412; O15301;  
DE 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE T-box transcription factor TBX5 (T-box protein 5).  
GN TBX5.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN [1]  
RP MEDLINE-97141914; PubMed-998164;  
RX LI O.V., Neubury-Ecob R., Terrett J.A., Wilson D.I., Curtis A.,  
RA Yi C.H., Bulleid P.J., Strachan T., Robson S., Bonnet D.,  
RA Young I.R., Raeburn J.A., Buckler A.J., Gebuhr T., Law D.J.,  
RT "Holt-Oram syndrome is caused by mutations in TBX5, a member of the  
RT Brachyury (T) gene family";  
RL Nat. Genet. 15:21-29(1997).  
[2]  
RP SEQUENCE FROM N.A. ALTERNATIVE SPLICING, AND VARIANT HOS GLN-237.  
RX MEDLINE-97141915; PubMed-998165;  
RA Basson C.T., Bachinsky D.R., Lin R.C., Levi T., Elkins J.A.,  
RA Soultis J., Grayzel D., Kroumpouzou E., Traill T.A.,  
RA Leblanc-Stracskai J., Renault B., Kucherlapati R., Seidman J.G.,  
RA Seidman C.E.;  
RT "Mutations in human TBX5 cause limb and cardiac malformation in  
RT Holt-Oram syndrome";  
RL Nat. Genet. 15:30-35(1997).  
[3]  
RP SEQUENCE FROM N.A.  
RA Cross S.J., Ching Y.H., Armstrong-Buisseret L., Spranger S.,  
RA Munnich A., Bonnet D., Armstrong M., Jonveaux P., Newbury-Ecob R.,  
RA Brook D.;  
RT "The mutation spectrum of Holt-Oram syndrome";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
[4]  
RP VARIANTS HOS ARG-80; GLN-237 AND TRP-237.  
RX MEDLINE-99178990; PubMed-10077612;  
RA Basson C.T., Huang T., Lin R.C., Bachinsky D.R., Weremowicz S.,  
RA Vaglio A., Bruzone R., Quadrelli R., Lerone M., Romeo G., Silengo M.,  
RA Pereira A., Krieger J., Mesquita S.F., Kamisago M., Morton C.C.,  
RA Pierpont M.E., Muller C.W., Seidman J.G., Seidman C.E.;  
RT "Different TBX5 interactions in heart and limb defined by Holt-Oram  
RT syndrome mutations";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:2919-2924(1999).  
CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES  
CC REQUIRED FOR MESODERM DIFFERENTIATION. PROBABLY PLAYS A ROLE IN  
CC LIMB PATTERN FORMATION.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

P79944 xenopus lae  
Q17134 branchiosteo  
P79777 gallus gall  
Q90117 homo sapien  
P80492 branchiosteo  
O60806 homo sapien  
Q25113 hemicentrot  
O95939 homo sapien  
O54839 mus musculu  
P79778 gallus gall  
P56158 halocynthia  
P55965 drosophila

#### ALIGNMENTS

34 532.5 28.2 692 1 EOMD\_XENLA  
35 531.5 28.1 448 1 BRAL\_BRAFL  
36 531 28.1 433 1 BRAC\_CHICK  
37 528.5 28.0 535 1 TX21\_HUMAN  
38 527.5 27.9 440 1 BRA2\_BRAFL  
39 513 27.2 448 1 TX19\_HUMAN  
40 510 27.0 434 1 BRAC\_HEMPU  
41 507 26.8 686 1 EOMD\_HUMAN  
42 505 26.7 688 1 EOMD\_MOUSE  
43 497 26.3 397 1 TBAT\_CHICK  
44 496.5 26.3 471 1 BRAC\_HALRO  
45 489 25.9 697 1 BYN\_DROME

[illegible]



```

RN RP SEQUENCE FROM N.A.
RX MEDLINE 17979242; Pubmed-9878690;
RA Aquilnik S.I., Garvey N., Hancock S., Ruvinisky I., Charman D.L.,
RT Aquilnik I., Bollag R.J., Papaioannou V.E., Silver L.M.;
RT "Evolution of mouse T-box genes by tandem duplication and cluster
RT dispersion.";
RL Genetics 144:249-254(1996).
RN RP [2]
RX MEDLINE 570664; Pubmed-9853987;
RA Chapman D.I., Garvey N., Hancock S., Alexiou M., Aquilnik S.I.,
RA Gibson-Brown J.J., Cebra-Thomson J., Bollag R.J., Silver L.M.,
RA Papaioannou V.E.;
RT "Expression of the T-box family genes, Tbx1-Tbx5, during early mouse
RT development.";
RL Dev. Dyn. 206:379-390(1996).
CC CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
CC CC LMB PATTERN FORMATION, DIFFERENTIATION. PROBABLY PLAYS A ROLE IN
CC CC LMB PATTERN FORMATION.
CC CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC CC THE ALLANTOIS WHERE EXPRESSION CONTINUES THROUGH DAY 8.5, AT DAY
CC CC 9.5, EXPRESSION IS FOUND IN THE GENITAL PAPILLA, BODY WALL AND
CC CC LMB BUDS (HIGHER LEVELS IN HINDLIMB). AT DAY 12.5, EXPRESSED IN
CC CC THE MESENCHYME OF THE MANDIBULAR ARCH, OF THE LONG AND OF THAT
CC CC SHORT BRANCHES OF THE MANDIBULAR ARCH.
CC CC -1- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC at the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL: U01279; AAC53108.1;
CC CC HSP: P24781; IYBR.
CC CC TRANSFAC: T04350; Tbx4.
CC CC InterPro: IPR001699; TF_T-box.
CC CC Pfam: PF00507; T-box; 1.
CC CC SMART: SM00425; TBOX_1.
CC CC PROSITE: PS01283; TBOX_1; 1.
CC CC PROSITE: PS01264; TBOX_2; 1.
CC CC PROSITE: PS0252; TBOX_3; 1.
CC CC Transcription regulation; DNA-binding; Nuclear protein;
CC CC Non-transcriptional protein;
CC CC DNA_BIND <1 164
CC CC NON_TER 173 173
CC CC SEQUENCE 173 AA; 19790 MW; 4A5238290E5B075D CRC64;
Query Match 45.8%; Score 865.5; DB 1; Length 173;
Best Local Similarity 93.6%; Pred. No. 9,5e-65;
Matches 160; Conservative 5; Mismatches 5; Indels 1; Gaps 1;
QY 74 MIITKAGRMFPSPYKVKVTLGNPKTKYILLMDIVPADDRHYKFADNKSVYTKAEPMP 133
DQ 1 MIITKAGRMFPSPYKVKVTLGNPKTKYILLMDIVPADDRHYKFADNKSVYTKAEPMP 59
QY 134 RLYVHPDSPATGAHMRQLVSPQKLITNNHLPFGHITILNSMKHYOPRLHYKADENNG 193
DQ 60 RLYVHPDSPATGAHMRQLVSPQKLITNNHLPFGHITILNSMKHYOPRLHYKADENNA 119
QY 194 FGSKNTAFCTHVPETAFIVTSYONHKITOLKIENNPFAKGFSGSDMEL 244
DQ 120 FGSKNTAFCTHVPETFSIVTSYONHKITOLKIENNPFAKGFSGSDSL 170
Query Match 34.9%; Score 658.5; DB 1; Length 423;
Best Local Similarity 45.3%; Pred. No. 4.2e-47;
Matches 144; Conservative 47; Mismatches 84; Indels 43; Gaps 9;
QY 30 SALGAPSKSPSPQAFT--QQQM-----EGI---KVFHERELMLKAFIEVGT 72
DQ 20 AGVGPGAGCGPPNLFPSMLOAGFPDGVGSPDDGVTDHDPKVELDEHLMQOFSQCT 79
QY 73 EMITTKAGRMFPSPYKVKVTLGNPKTKYILLMDIVPADDRHYKFADNKSVYTKAEPMP 132
DQ 80 EWITKSGRIFPAPYKVKISGLDKKSOYFVMMDLVPADDRHYKFNNSRMNIAKADPEMP 139
QY 133 GRLYVHPDSPATGAHMRQLVSPQKLITNNHLPFGHITILNSMKHYOPRLHYK-ADEN 191
DQ 140 KFLYVHPDSPATGAHMRQLVSPQKLITNNHLPFGHITILNSMKHYOPRLHYKADENH 199
QY 192 NCGFSKNTAFCTHVPETAFIVTSYONHKITOLKIENNPFAKGFSGS-----DQKEL 245
DQ 200 NLMS--TFTFVFRETETFIATVQNEKVTLEKIENNPFAKGFSGKREKROLH 256
QY 246 RMRKQSKYVPVPRSTVROKVAS--NRPFSSESRALESNNL-----GSOYCENGV 297
DQ 257 RMNGDQTSQPP-----GRYASLPTPHSPSESNSEDEPTLKKCKPEPSPQPTSSSL 308
QY 298 SGPSQDLPPPNYPVLPQ 315
DQ -309 STSTPTTLSAHHPLRSPQ 326

```



```

OY 17 PDKNDLPCDSKPSALGAKPSGSSPOAAFTQOGMEGKVFLEHRELWLKTFHEVTEM11 76
DB 5 PDIR APCD-----KASFLDEP-YFSSIT-----VLEEDGLMKFRQIGTEM11 49
OY 77 TACGRMPFSYKVKYGLNPKTKYLLMDIVPADDRHYKADNKKSVTKCAEPAMPGLY 136
DB 50 TKSGRMPFPOCKIKVSGLIPEYAKYLMVLDFVPVDNFRYKKNKDQMEVAGKAEPOLPCRTY 109
OY 137 VHPDSPATCAHMRQVNSFOKLKTNHHLDPFGHILNSMHKYOPRLHIVTKADENNGFS 196
DB 110 VNPDSFCAHMRQVNSFOKLKTNHHLDPFGHILNSMHKYOPRLHIVTKADENNGFS--FSV 167
OY 197 KNTAFCTHVPETAPJAVTSYONKHITKLTENNPAGKER-----GSDMELHRM 247
DB 168 RNSIFQVFSFETVFTSVAYNQEDITLKLIDNNPKAFKGFREHNGKTRREGRAKSKSP 227
OY 248 SMOKSEYVPVPASTVROK-----VASNHSFP-----SSSRALSTSN 286
DB 228 AGQRKLPEKSGNERDEPKEDNDVVKESNPVYSGGTFPPISEONSHAFPAASP 287
OY 287 LGSQYOCENGSGPSQDLPPNPVLPQEHUS-----QIYHCTKRGCECHP 333
DB 288 APABOR-----EGPAREQOVPPSYOTYRFEAGDSQOLPSRDVAALNDFGRCH-HP 338

RESULT 8
TXB6_HUMAN
ID TXB6_HUMAN STANDARD; PRT; 701 AA.
AC Q60707;
DT 01-NOV-1997 (Rel. 35, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE T-box transcription factor TBX2 (T-box protein 2).
GN TBX2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_Taxid=10090;
RN [1]
SEQUENCE FROM N.A., AND DEVELOPMENTAL EXPRESSION.
RP TISSUE=Embryo;
RC MEDLINE=95004605; PubMed=7920656;
RA Bollag R.J., Siegfried Z., Cebra-Thomas J.A., Garvey N., Davison E.M.,
RT 'an ancient family of embryonically expressed mouse genes sharing a
RL conserved protein motif with the T locus.';
RN [2]
SEQUENCE FROM N.A.
RP PubMed=10770922;
RA Ferreira S., Liu B., Goding C.R.;
RT 'the T-box transcription factor Tbx2 is a target for the
RL microphthalmia-associated transcription factor in melanocytes.';
RN [3]
DEVELOPMENTAL EXPRESSION.
RA Chapman D.L., Garvey N., Hancock S., Alexiou M., Aguilnik S.I.,
RA Gibson-Brown J.J., Cebra-Thomas J., Bollag R.J., Silver L.M.,
RT 'Expression of the T-box family genes, Tbx1-Tbx5, during early mouse
RL development.';
RN [4]
DEV. Dyn. 206:379-390(1996).
CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
CC REQUIRED FOR MESODERM DIFFERENTIATION. PROBABLY PLAYS A ROLE IN
CC LIMB PATTERN FORMATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HIGHEST LEVELS IN LUNG. ALSO FOUND
CC IN HEART, KIDNEY AND OVARY.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION FIRST OBSERVED AT DAY 9.5 IN
CC THE OTIC AND OPTIC VESICLES AND IN THE FACIAL REGION. AT DAY
CC 12.5, EXPRESSED IN THE TRIGEMINAL GANGLIA, FACIAL REGIONS, RETINA

```

```

CC AND LIMB BUD MESENCHYME. IN LATER STAGES, FOUND IN EAR PINNAE,
CC THE MILK LINE, LUNG MESENCHYME, BODY WALL, GENITAL RIDGE AND
CC DEVELOPING NERVOUS SYSTEM.
CC -1- SIMILARITY: CONTAINS T-box DOMAIN
CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U155566; AAC52697.1;
CC EMBL: AF244917; AAF90050.1;
CC HSP: P24781; 1XBR.
CC TRANSFAC: T04358;
CC MGD; MGI:98494; Tbx2.
CC InterPro: IPR001699; T-box.
CC Pfam: PF00907; T-box.
CC SMART: SM00427; T-box.
CC PROSITE: PS01283; TBOX_1; 1.
CC PROSITE: PS01264; TBOX_2; 1.
CC PROSITE: PS0252; TBOX_3; 1.
CC Transcription regulation; DNA-binding; Nuclear protein;
CC Developmental protein.
CC DOMAIN 48 53 POLY-A.
CC TRANSFAC: T04358;
CC DOMAIN 572 580 POLY-A.
CC DOMAIN 586 594 POLY-A.
CC SEQUENCE 701 AA; 74244 MW; 8D90ED6DA32B3859 CRC64;
SQ
Query Match 34.4%; Score 649.5; DB 1; Length 701;
Best Local Similarity 44.3%; Pred. No. 4.6e-46;
Matches 145; Conservative 41; Mismatches 82; Indels 59; Gaps 8;
OY 23 PCDSKPSALGAKPSGSSPOAAFTQOGMEGKVFLEHRELWLKTFHEVTEM11 51
DB 29 PALALPPGALGKPLDPDGLAGAAAAAAGHVSALGPPPAHLRLSKLSLEPE 88
OY 52 ----EGIKVFLHRELWLKTFHEVTEM11TKAGRMFPFSYKVKYGLNPKTKYLLMDIV 107
DB 89 DEVEDDPKVTLEAKELMDWDFHKLGTWYTKSGRRMFPFKVRSGLDKYKYLMDIV 148
OY 108 PADDRHYKADNKKSVTKCAEPAMPGLYVHPDSPATCAHMRQVNSFOKLKTNHLDP 167
DB 149 AADCRVKFHSRWVASKADPEMKRWYIHPDSPATGEQMAKPVAFHKLKLTNNISDK 208
OY 168 FGHILNSMHKYOPRLHIVTKADENNGFSKNTAFCTHVPETAPJAVTSYONKHITOLKI 227
DB 209 HGFTILNSMHKYOPRLHIVRA--NDILKLPYSTERTYVPPETDFAVTAYQNDKITOLKI 266
OY 228 ENNPANGSGRSGDMDLHRSNOSKEYPVVPRSTVROKVASHSFSSSRALSTSNUL 287
DB 267 DNNPFAKGFRTGNG-----CRKKRKLTLDP--TLR--LYVEHCKPREGAEDASS-- 314
OY 288 GSQYQCEGVSGPSQDLPPNPVLP 314
DB 315 -----CD-----PPAREPPPSAAP 331
RESULT 9
TXB6_HUMAN
ID TXB6_HUMAN STANDARD; PRT; 436 AA.
AC O85947;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
GN TBX6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```







[illegible]





```

Query Match      32.5%; Score 616.5; DB 1; Length 251;
Best Local Similarity 53.5%; Pred No. 6.2e-44;
Matches 122: Conservative 34; Mismatches 55; Indels 17; Gaps 4;

QY 33 GAPSKSPSPQAA-----FTQGMGKIVFLHIERELWLKFEVGTGEMIIIRAGR 82
DB 24 GSGSGSPSSSICTEPIPTPIPSBMAKACSLTELKELMDNFHELGTEMIIIRSGR 83
QY 83 MFPSYKVKYGLNPKTKYLLIMDIVPADDHRYKADNK--WSVTGKAEPAMGRGLVYHPD 140
DB 84 MFTIRVSFSGVDEAKYIVLMDIVPDKNRYAYHRSSWLKAGKADPLPARLYHPD 143
QY 141 SPATCAHAWROLVVSOKLKTNNHLDPGHIILNSMKHYOPRLHTVKADENNG--FGSKN 198
DB 144 SPFTGEOLLKQVSEKVKLTNNELDQGHILLNSMKHYOPRVHIIKKDKHTASILLKAS 203
QY 199 TAFCTHVFPETAFTAVTSYQNHKITYOLKIENPFAGFRGSD--DME 243
DB 204 EEFRTIFFEIVFTVTAIQNOLITRLKIDSHPFAGFRDSSRLADIE 251

```

Search completed: June 17, 2003, 09:47:36  
Job time : 13 secs



Tue Jun 17 10:11:35 2003

us-10-027-059A-1.rspt

Page 2

```

Oy 61 RELMLKFEHVGTEMLITKACGRMPSPSKYKYNGLNPKTKYIILMDIYVADPHRYKFAONK 120
Db 78 RELMLKFEHVGTEMLITKACGRMPSPSKYKYNGLNPKTKYIILMDIYVADPHRYKFAONK 137
Oy 121 WSYTGKAEPAMPGLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPEGHILNLSMHKYO 180
Db 138 WSYTGKAEPAMPGLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPEGHILNLSMHKYO 197
Oy 181 PRHIVKADENNGFGSKNTACHTVPEPTAFIAYTSYONKHITOLKLTENNPRAKFGPSD 240
Db 198 PRHIVKADENNGFGSKNTACHTVPEPTAFIAYTSYONKHITOLKLTENNPRAKFGPSD 257
Oy 241 DMLHRMSRMOSKEEYVPRSTYRQKVASNSHSPSESSEALSTSSNLGSOYCENGVSOP 300
Db 258 DMLHRMSRMOSKEEYVPRSTYRQKVASNSHSPSESSEALSTSSNLGSOYCENGVSOP 317
Oy 301 SODLLPPNPPYPLPOEHSOIYHCTKRKECSTTEHRY 334
Db 318 SODLLPPNPPYPLPOEHSOIYHCTKRKECSTTEHRY 355

```

RESULT 2

```

Oy 09PM68 PRELIMINARY: PRT: 521 AA.
AD 09PM68:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE T-box transcription factor Tbx5.
OS Gallus gallus (Chicken).
OC Archaeopteryx; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Actinopterygii; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
NCI_Taxid=9031;
RN 11
SEQUENCE FROM N.A.
RA Bruneau B.G., Logan M., Davis N., Levi T., Tablin C.J., Seidman J.G.,
RT Selman C.E.;
RT PubMed-99303540; PubMed-10373308;
RT Human transcription factor Tbx5 mRNA, alternatively spliced.
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB069396; AAC23683.2;
DR HSSR: Z24781; 118R.
DR DDB: J039198.1; 118R.
DR InterPro: IPR001699; TF-T-box.
DR Pfam: PF00907; T-box_1.
DR PRINTS: PRO0937; TBOX.
DR SMART: SM0423; TBOX_1.
DR PROSITE: PS01283; TBOX_1.
DR PROSITE: PS01264; TBOX_2.
DR PROSITE: PS0252; TBOX_3.
DR PROSITE: PS0252; TBOX_3.
SO SEQUENCE 521 AA: 58402 MW: 17633DOFIDGCE308 CRC64;

```

Query Match 85.3%; Score 1611; DB 13; Length 521;  
 Best Local Similarity 88.8%; Pred. No. 3 de-138;  
 Matches 300; Conservative 15; Mismatches 19; Indels 4; Gaps 1;

```

Oy 1 MDADEGFGIARFLEPPARDLPCSCPSALGAPSPSPSPALFQOGMGKGFVLE 60
Db 1 MDADEGFGIARFLEPPARDLPCSCPSALGAPSPSPSPALFQOGMGKGFVLE 60
Oy 61 RELMLKFEHVGTEMLITKACGRMPSPSKYKYNGLNPKTKYIILMDIYVADPHRYKFAONK 120
Db 61 RELMLKFEHVGTEMLITKACGRMPSPSKYKYNGLNPKTKYIILMDIYVADPHRYKFAONK 120
Oy 121 WSYTGKAEPAMPGLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPEGHILNLSMHKYO 180
Db 121 WSYTGKAEPAMPGLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPEGHILNLSMHKYO 180

```

```

Db 121 WSYTGKAEPAMPGLYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPEGHILNLSMHKYO 180
Oy 181 PRHIVKADENNGFGSKNTACHTVPEPTAFIAYTSYONKHITOLKLTENNPRAKFGPSD 240
Db 181 PRHIVKADENNGFGSKNTACHTVPEPTAFIAYTSYONKHITOLKLTENNPRAKFGPSD 240
Oy 241 DMLHRMSRMOSKEEYVPRSTYRQKVASNSHSPSESSEALSTSSNLGSOYCENGVSOP 300
Db 241 DMLHRMSRMOSKEEYVPRSTYRQKVASNSHSPSESSEALSTSSNLGSOYCENGVSOP 300
Oy 301 SODLLPPNPPYPLPOEHSOIYHCTKRKECSTTEHRY 334
Db 301 SODLLPPNPPYPLPOEHSOIYHCTKRKECSTTEHRY 338

```

RESULT 3

```

Oy 09GT80 PRELIMINARY: PRT: 342 AA.
AD 09GT80:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Transcription factor T-box 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Actinopterygii; Primates; Catarrhini; Hominoidea; Homo.
NCI_Taxid=9606;
RN 11
SEQUENCE FROM N.A.
RA Denuva M., Yoneda T., Takeda Y., Furukawa K., Mabuchi H.;
RT Human transcription factor Tbx5 mRNA, alternatively spliced.
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB069396; AAC23683.2;
DR HSSR: Z24781; 118R.
DR DDB: J039198.1; 118R.
DR InterPro: IPR001699; TF-T-box.
DR Pfam: PF00907; T-box_1.
DR PRINTS: PRO0937; TBOX.
DR PROSITE: PS01283; TBOX_1; UNKNOWN.1.
DR PROSITE: PS01264; TBOX_2; UNKNOWN.1.
DR PROSITE: PS0252; TBOX_3; 1.
DR PROSITE: PS0252; TBOX_3; 1.
SO SEQUENCE 342 AA: 38916 MW: C368AD932DEPDA CRC64;

```

Query Match 79.6%; Score 1503; DB 4; Length 342;  
 Best Local Similarity 97.6%; Pred. No. 1 de-128;  
 Matches 281; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

```

Oy 51 MEDIVFLEHRELMLKFEHVGTEMLITKACGRMPSPSKYKYNGLNPKTKYIILMDIYVADPH 110
Db 1 MEDIVFLEHRELMLKFEHVGTEMLITKACGRMPSPSKYKYNGLNPKTKYIILMDIYVADPH 60
Oy 111 DHRVTFADKNSYVGAEPAMGRILYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPEGH 170
Db 61 DHRVTFADKNSYVGAEPAMGRILYVHPDSPATGAMHMRQVLSFOKLKLTNNHLDPEGH 120
Oy 171 IILNSMHKYOPLHIVKADENNGFGSKNTACHTVPEPTAFIAYTSYONKHITOLKLTENN 230
Db 121 IILNSMHKYOPLHIVKADENNGFGSKNTACHTVPEPTAFIAYTSYONKHITOLKLTENN 180
Oy 231 PRKAFGSDMLHRMSRMOSKEEYVPRSTYRQKVASNSHSPSESSEALSTSSNLGSO 290
Db 181 PRKAFGSDMLHRMSRMOSKEEYVPRSTYRQKVASNSHSPSESSEALSTSSNLGSO 240
Oy 291 YOCENGVSOPDILLPPNPPYPLPOEHSOIYHCTKRKECSTTEHRY 334
Db 241 YOCENGVSOPDILLPPNPPYPLPOEHSOIYHCTKRKECSTTEHRY 288

```

RESULT 4

```

Oy 09W7C2 PRELIMINARY: PRT: 519 AA.
AD 09W7C2:
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

```



DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE T-box transcription factor tbox5.  
 GN tbox5.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenoipidae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99180578; PubMed=10079235;  
 RA Horb M.E., Thomson G.H.;  
 RT "tbox5 is essential for heart development.";  
 RL Development 126:1739-1751(1999).  
 DR EMBL: AF133036; AACD3592.1; -;  
 DR HSSP: P24781; 1XBR.  
 DR TRANSFAC: T04422; -;  
 DR Interpro: IPR001699; TF-T-box.  
 DR Pfam: PF00907; T-box; 1.  
 DR PRINTS: PR00937; TBOX.  
 DR SMART: SM00425; TBOX. 1.  
 DR PROSITE: PS01283; TBOX\_1; 1.  
 DR PROSITE: PS01264; TBOX\_2; 1.  
 DR PROSITE: PS01252; TBOX\_3; 1.  
 SO SEQUENCE 519 AA; 56080 MW; 8B5E096C67C3D1E CRC64;

Query Match 79.5%; Score 1501.5; DB 13; Length 519;  
 Best Local Similarity 86.0%; Pred. No. 3, 3e-126;  
 Matches 262; Conservative 16; Mismatches 29; Indels 1; Gaps 1;

QY 1 MADADGGLAHTPLPPADKLDPCSKRESALGAPSKSPSSPOAAFTQOGMGIGKIVFLHE 60  
 DB 1 MADTEAAGMDPTFAEPKELCEKDDQJASAKSKPTSPAPFTQOGMGIGKIVFLHE 60  
 QY 61 RELMLKFHEVGTETMITTAGRRMFPYKYVYVGLMPKTKYLLMDIVPADDHRRKFPADNK 120  
 DB 61 RELMLKFHEVGTETMITTAGRRMFPYKYVYVGLMPKTKYLLMDIVPADDHRRKFPADNK 120  
 QY 121 WSVTGAEPAMPGRGLYHPDSPATGAHMQQLYSOKLKTJNNHLDPFCHILLNSMKRYO 180  
 DB 121 WSVTGAEPAMPGRGLYHPDSPATGAHMQQLYSOKLKTJNNHLDPFCHILLNSMKRYO 180  
 QY 181 PRLHYKADENNGFGSKNTAFCTHVPETAFATVTSYONKKTITOLKTENNPPAGFRGSD 240  
 DB 181 PRLHYKADENNGFGSKNTAFCTHVPETAFATVTSYONKKTITOLKTENNPPAGFRGSD 240  
 QY 241 DMLHHRMSMOSKEYPVYPRSTYVROKVASNHSFSSRALSTSNLGSQYCCNCGVSGP 300  
 DB 241 DMLHHRMSMOSKEYPVYPRSTYVROKVASNHSFSSRALSTSNLGSQYCCNCGVSGP 300  
 QY 301 SQDLPLPPNRYPLPQEHISQIYCTKKR 327  
 DB 301 SQDLPLPPNRYPLPQEHISQIYCTKKR 327  
 QY 301 SQDLPLPPNRYPLPQEHISQIYCTKKR 327  
 DB 301 SQDLPLPPNRYPLPQEHISQIYCTKKR 327

## RESULT 5

Q9PUS7 PRELIMINARY; PRT; 485 AA;  
 AC Q9PUS7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 GN T-box transcription factor tbox5.  
 GN tbox5.  
 OS Brachydanio rerio (zebrafish) (zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 CC Cyprinidae; Danio.  
 NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9425182; PubMed=10495283;  
 RA Tamura K., Toney-Tamura S., Belmonte J.C.;

RT "Differential expression of tbox4 and tbox5 in zebrafish fin buds."  
 RL Mech. Dev. 87:181-184(1999).  
 DR EMBL: AF152607; AAF06733.1; -;  
 DR HSSP: P24781; 1XBR.  
 DR ZFIN: ZDB-GENE-991124-7; tbox5.  
 DR Interpro: IPR001699; TF-T-box.  
 DR Pfam: PF00907; T-box; 1.  
 DR PRINTS: PR00937; TBOX.  
 DR SMART: SM00425; TBOX. 1.  
 DR PROSITE: PS01283; TBOX\_1; 1.  
 DR PROSITE: PS01264; TBOX\_2; 1.  
 DR PROSITE: PS01252; TBOX\_3; 1.  
 SO SEQUENCE 485 AA; 54322 MW; E969A859C8D7B791 CRC64;

Query Match 73.7%; Score 1392; DB 13; Length 485;  
 Best Local Similarity 78.5%; Pred. No. 3e-118;  
 Matches 267; Conservative 24; Mismatches 39; Indels 10; Gaps 5;

QY 1 MADADGGLAHTPLPPADKLDPCSKRESALGAPSKSPSSPOAAFTQOGMGIGKIVFLHE 60  
 DB 1 MADTEAAGMDPTFAEPKELCEKDDQJASAKSKPTSPAPFTQOGMGIGKIVFLHE 60  
 QY 61 RELMLKFHEVGTETMITTAGRRMFPYKYVYVGLMPKTKYLLMDIVPADDHRRKFPADNK 120  
 DB 61 RELMLKFHEVGTETMITTAGRRMFPYKYVYVGLMPKTKYLLMDIVPADDHRRKFPADNK 119  
 QY 121 WSVTGAEPAMPGRGLYHPDSPATGAHMQQLYSOKLKTJNNHLDPFCHILLNSMKRYO 180  
 DB 121 WSVTGAEPAMPGRGLYHPDSPATGAHMQQLYSOKLKTJNNHLDPFCHILLNSMKRYO 179  
 QY 181 PRLHYKADENNGFGSKNTAFCTHVPETAFATVTSYONKKTITOLKTENNPPAGFRGSD 240  
 DB 181 PRLHYKADENNGFGSKNTAFCTHVPETAFATVTSYONKKTITOLKTENNPPAGFRGSD 239  
 QY 241 DMLHHRMSMOSKEYPVYPRSTYVROKVASNHSFSSRALSTSNLGSQYCCNCGVSGP 299  
 DB 241 DMLHHRMSMOSKEYPVYPRSTYVROKVASNHSFSSRALSTSNLGSQYCCNCGVSGP 299  
 QY 300 SQDLPLPPNRYPLPQEHISQIYCTKKR--GEC--DHPW 334  
 DB 300 TSQDLPLPPNRYPLPQEHISQIYCTKKR--GEC--DHPW 336

## RESULT 6

Q9TAK8 PRELIMINARY; PRT; 492 AA;  
 AC Q9TAK8;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 GN T-box transcription factor tbox5.  
 GN tbox5.  
 OS Brachydanio rerio (zebrafish) (zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 CC Cyprinidae; Danio.  
 NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20130955; PubMed=10664151;  
 RA Ruvinsky I., Oates A.C., Silver L.M., Ho R.K.;  
 RT "The evolution of paired appendages in vertebrates: T-box genes in the zebrafish."  
 RL Dev. Genes Evol. 210:87-91(2000).  
 DR EMBL: AF179407; AAF59837.1; -;  
 DR HSSP: P24781; 1XBR.  
 DR Interpro: IPR001899; Gram\_pos\_anchor.  
 DR Interpro: IPR001699; TF-T-box.  
 DR Pfam: PF00907; T-box; 1.  
 DR PRINTS: PR00937; TBOX.  
 DR SMART: SM00425; TBOX. 1.  
 DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; UNKNOWN.1.  
 DR PROSITE: PS01283; TBOX\_1; 1.  
 DR PROSITE: PS01264; TBOX\_2; 1.

DR PROSITE: PSS0252; TBOX\_3; 1.  
 SO SEQUENCE 492 AA; 55236 MW; 566940D1942D2750 CRC64;  
 Query Match 73.4%; Score 1387; DB 13; Length 492;  
 Best Local Similarity 76.2%; Pred. No. 8, 6e-118;  
 Matches 266; Conservative 24; Mismatches 40; Indels 10; Gaps 5;

1 MADADGFGALTFLEPDADKLDPCDSKESALGAPSKSPSSPOAFTQOCMEGICVLEH 60  
 DB 1 MASSEDFRLQNSDSBEPKLDNCKSDONMAVSNSSS-QTYTGGHGLVYVLE 59  
 OY 61 RELMLKPFHEVGTEMLITKAGRRFPSSYKVVYGLNPTPTLYLLMDIVPADRHRYKFA 120  
 DB 60 RELMLKPFHEVGTEMLITKAGRRFPSPKVPYGLNPTPTLYLLMDIVPADRHRYKFA 119  
 OY 121 MASTKAPKMGRLVHPDSFGAHMROLYSPKLTNNHLDPPGHITLNSMKHYO 180  
 DB 120 MASTKAPKMGRLVHPDSFGAHMROLYSPKLTNNHLDPPGHITLNSMKHYO 179  
 OY 181 PRLATVADENNGFGSKNTACTHVPETAFIATVSYONHKTITQLEINFPKQFGSD 240  
 DB 180 PRLATVADENNGFGSKNTACTHVPETAFIATVSYONHKTITQLEINFPKQFGSD 239  
 OY 241 DMLHARMSRMOS-KEYVVPVPSYVROKAVNSPSSSESRALSTSNIGSGOCENGCVG 299  
 DB 240 DMLHARMSRMOSTKEYVVPVPSYVROKAVNSPSSSESRALSTSNIGSGOCENGCVG 299  
 OY 300 PPSDILPPNPVPLPDHNSQIYCTCKRK--GEC--DHPM 334  
 DB 300 TSDQILLQSSSY--HEHTQDYHCIRKRVEDCPCAGHRY 336

RESULT 7  
 ID PROSITE: PSS0252; TBOX\_3; 1.  
 AC O93TK3 PRELIMINARY; PRT; 372 AA.  
 DT 01-MAY-2000 (TREMblrel). 13, Created  
 DT 01-MAY-2000 (TREMblrel). 13, Last sequence update  
 DT 01-MAR-2002 (TREMblrel). 20, Last annotation update  
 DR T-box containing transcription factor Tbx5.1 (Fragment).  
 OS Brachydanio rerio (zebrafish) (zebra danio).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 RN NCBI\_TaxID:7955;  
 RX MEDLINE:20108594; PubMed:10640716;  
 RA Begemann G., Ingham P.W.,  
 Mech. Dev. 90:299-304(2000).  
 DR HSSP: P24781; AA02296.1;  
 DR ZFIN: ZDB-GENE-991124-7; Tbx5.  
 DR InterPro: IPR001699; TF-T-box.  
 DR Pfam: PF00907; T-box\_1.  
 DR PRINTS: PR00937; TBOX.  
 DR SMART: SM00423; TBOX\_1;  
 DR PROSITE: PS01283; TBOX\_1;  
 DR PROSITE: PS01284; TBOX\_2;  
 DR PROSITE: PSS0252; TBOX\_3; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 372 372  
 SO SEQUENCE 372 AA; 41764 MW; 6DF0DFD3C5340BC CRC64;  
 Query Match 71.4%; Score 1349; DB 13; Length 372;  
 Best Local Similarity 84.0%; Pred. No. 7, 2e-111;  
 Matches 258; Conservative 15; Mismatches 24; Indels 10; Gaps 5;

34 APSSKSPSSPOAFTQOCMEGICVLEHRELMLKPFHEVGTEMLITKAGRRFPSSYKVVYGLN 93  
 DB 4 AYSSKSPSS-OTTYTGGHGLVYVLEHRELMLKPFHEVGTEMLITKAGRRFPSSYKVVYGLN 62

OY 94 LMRKTYLLMDIVPADRHRYKFAFNKNSYTGKAPKAPGRLYVHPDSFGAHMROLY 153  
 DB 63 LMRKTYLLMDIVPADRHRYKFAFNKNSYTGKAPKAPGRLYVHPDSFGAHMROLY 122  
 OY 154 SFQKLTNNHLDPPGHITLNSMKHYORLHYKADENNGFGSKNTAFCTHVPETAFIA 213  
 DB 123 SFQKLTNNHLDPPGHITLNSMKHYORLHYKADENNGFGSKNTAFCTHVPETAFIA 182  
 OY 214 VTSYONHKTITQLEINFPKQFGSDMEIHNRSQSG-KEYVVPVPSYVROKAVNS 272  
 DB 183 VTSYONHKTITQLEINFPKQFGSDMEIHNRSQSG-KEYVVPVPSYVROKAVNS 242  
 OY 273 PFSSSESRALSTSNIGSGOCENGCVGSDLLPPNPVPLPDHNSQIYCTCKRK--GEC 330  
 DB 243 PFSSSESRALSTSNIGSGOCENGCVGSDLLPPNPVPLPDHNSQIYCTCKRK--GEC 299  
 OY 331 --DHPM 334  
 DB 300 PACGHPY 306

RESULT 8  
 ID PROSITE: PSS0252; TBOX\_3; 1.  
 AC O93288 PRELIMINARY; PRT; 541 AA.  
 DT 01-NOV-1998 (TREMblrel). 08, Created  
 DT 01-NOV-1998 (TREMblrel). 08, Last sequence update  
 DT 01-MAR-2002 (TREMblrel). 20, Last annotation update  
 DR T-box Protein 4 (T-box transcription factor Tbx4).  
 OS Gallus gallus (Chicken).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 RN NCBI\_TaxID:9031;  
 RX MEDLINE:98220375; PubMed:9550719;  
 RA Issac K., Rodiguez-Esteban C., Ryan A., Altshuler M., Tsukui T.,  
 Patel K., Tickle C., Izpisua-Belmonte J. C.,  
 Development 125:1867-1875(1998).  
 RN 121  
 RP SEQUENCE OF 66-249 FROM N.A.  
 RA Logan M., Simon H.-G., Tablin C.,  
 RT Differential regulation of T-box transcription factors suggests a  
 DR T-box containing transcription factor Tbx4.  
 DR EMBL: AF033670; AAC41298.1;  
 DR EMBL: AF069395; AAC23682.1;  
 DR HSSP: P24781; Tbx4.  
 DR InterPro: IPR001699; TF-T-box.  
 DR Pfam: PF00907; T-box\_1.  
 DR SMART: SM00425; TBOX\_1.  
 DR PROSITE: PS01283; TBOX\_1;  
 DR PROSITE: PS01284; TBOX\_2;  
 DR PROSITE: PSS0252; TBOX\_3; 1.  
 SO SEQUENCE 541 AA; 60381 MW; 5744194AB7CC22EB CRC64;  
 Query Match 61.0%; Score 1152.5; DB 13; Length 541;  
 Best Local Similarity 64.4%; Pred. No. 2, 2e-96;  
 Matches 230; Conservative 33; Mismatches 79; Indels 15; Gaps 6;

1 MADADGFGALTFLEPDADKLDPCDSKESALGAPSKSP-SSPOAFTQOCMEGICVLEH 59  
 DB 7 LMRKTYLLMDIVPADRHRYKFAFNKNSYTGKAPKAPGRLYVHPDSFGAHMROLY 122  
 OY 60 RELMLKPFHEVGTEMLITKAGRRFPSSYKVVYGLNPTPTLYLLMDIVPADRHRYKFA 119  
 DB 63 EKLMLKPFHEVGTEMLITKAGRRFPSSYKVVYGLNPTPTLYLLMDIVPADRHRYKFA 122



Accession	Gene	Species	Length	Score	DB	Gap
RP	SEQUENCE FROM N.A.					
RP	MEDLINE=97236482; PubMed=9118806;					
RA	Simon H.G., Kittappa R., Khan P.A., Telford C., Liverange R.A.,					
RA	"Preliminary 11 of T-box genes in urodele amphibian limb development					
RT	and regeneration: candidate genes involved in vertebrate					
RT	forelimb/hindlimb patterning";					
RL	Development; 124:135-136(1997).					
RL	EMBL; U64433; AAB6743.1; -					
DR	HSSP; T24781; NBR					
DR	Plan; P001907; T-box; 1.					
DR	PRINTS; P000937; TBOX.					
DR	SMART; SMO0425; TBOX.1.					
DR	PROSITE; PS01564; TBOX.2; 1.					
DR	PROSITE; PS0252; TBOX.3; 1.					
SO	SEQUENCE 411 AA; 46092 MW; 77FB25B0077E675 CCK64;					
Query Match	57.5%; Score 1085.5; DB 13; Length 411;					
Best Local Similarity	87.0%; Pctd. No. 1.9e-90;					
Matches	7; Conservative 17; Idents 7; Gaps 4					
Oy	108 PADMRRKPKMKRSTGKAPKMGKLVHSDGATGAKHNNKQVSPQKTLNNHDP	167				
Db	1 PDDRRKPKFDMKSTGAPAPMGRLVYHDSPTAGAHNNQLVSFKLTNNHDP	160				
Oy	168 FGHILLSMHKQVRLHLYKADENKGGSNACTHPEVETAFVATVSTONKHITLKI	227				
Db	61 FGHILLSMHKQVRLHLYKADENKGGSNACTHPEVETAFVATVSTONKHITLKI	120				
Oy	228 ENNPAPGFGSDDELILHNSRMOSKEYPVVRSRTQKVASMHSPPSSALSTSNL	287				
Db	121 ENNPAPGFGSDDELILHNSRMOSKEYPVVRSRTQKVASMHSPPSSALSTSNL	179				
Oy	268 GSGYGCENYSGPSGDLIPPNRY-PIPOHSGYLYCKRKG--SC--DHNSTCL	339				
Db	180 SSQYCCENYSGPSGDLIPPNRY-PIPOHSGYLYCKRKG--SC--DHNSTCL	237				
RESULT 12						
ID	091BC5; PRELIMINARY; PRT; 133 AA.					
AC	01-OCT-2000 (TREMBL) 15, Created					
DT	01-OCT-2000 (TREMBL) 15, Last sequence update					
DT	01-MAR-2002 (TREMBL) 20, Last annotation update					
GN	T-box transcription factor Tbx3 (fragment).					
OS	Xenopus laevis (African clawed frog).					
OC	Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;					
OC	Xenopus; Xenopus.					
NC	NCBI_TaxId=8355;					
RP	SEQUENCE FROM N.A.					
RP	MEDLINE=2011077; PubMed=10704879;					
RA	Takabatake Y., Takabatake T., Takashima K.;					
RA	"Conserved and divergent expression of T-box genes Tbx2-Tbx5 in					
RT	Xenopus";					
RT	Development; 91:433-437(2000).					
RL	EMBL; AB012844; BAA93084.1; -					
DR	HSSP; P24781; 1XBR					
DR	TRANSFAC; T04423; -					
DR	InterPro; IPR001699; T-box; 1.					
DR	PRINTS; PR00937; TBOX.					
DR	PRINTS; PR00937; TBOX.					
DR	PROSITE; PS01564; TBOX.2; 1.					
DR	PROSITE; PS0252; TBOX.3; 1.					
FT	NON_TER 1 133					
SO	SEQUENCE 133 AA; 15209 MW; 9F5A287F9400BAEA CCK64;					

Query Match	Similarity	37.7%	Score 712	DB 13	Length 133	
Best Local	Similarity	98.5%	Pred. No. 4e-57			
Matches 131	Conservative	0	Mismatches 2	Indels 0	Gaps 0	
Db	1	DNKXSVTCGAAPMPCRLYVHPDSPATGACNMNOVYSPQKLTLTNNHLDPRCHILNSM	177			
Qy	118	DNKXSVTCGAAPMPCRLYVHPDSPATGACNMNOVYSPQKLTLTNNHLDPRCHILNSM	177			
Db	1	DKKNSVTCGAAPMPCRLYVHPDSPATGACNMNOVYSPQKLTLTNNHLDPRCHILNSM	177			
Qy	178	KYQPLHLVYKADENNGFSCNKAFCCTVHPPEFAIVYVSYONKKTOLKLTENNPFAGKR	237			
Db	61	KYQPLHLVYKADENNGFSCNKAFCCTVHPPEFAIVYVSYONKKTOLKLTENNPFAGKR	120			
Qy	238	GSDMDLHPNSRM 230				
Db	121	GSDMDLHPNSRM 133				
RESULT 13						
AC	09G0E7	PRELIMINARY	PRT:	174 AA.		
DT	01-MAR-2001	(TRENDELrel. 16, Created)				
DT	01-MAR-2001	(TRENDELrel. 15, Last sequence update)				
DT	01-JUN-2002	(TRENDELrel. 21, Last annotation update)				
DE	T-box protein Amphitritox/26 (fragment)					
OC	Eukaryota: Metazoa: Chordata: Cephalochordata: Branchiostomidae:					
NCBI	TextID-7739					
SEQUENCE FROM N.A.						
REDELrel-2019456	published-11033659					
Phylogenetic analysis of T-box genes demonstrates the importance of						
"genetics 156:1249-1257(2000) :						
EMBL: AF262565; AAC34890.1 :						
HSSP: P24781; I188						
RefSeq: P00057; T-box						
PRINTS: P000937; TBOX						
SMART: SM00425; TBOX_1						
PROSITE: PS01283; TBOX_1						
PROSITE: PS01264; TBOX_2						
PROSITE: PS50252; TBOX_3						
NON_TER	174					
SEQUENCE	174 AA: 20115 MW: EF36AC31678E880 CMC64					
Query Match						
Best Local	Similarity	75.7%	Score 674	DB 5	Length 174	
Matches 123	Conservative	24	Mismatches 26	Indels 2	Gaps 1	
Qy	63	LKTKPHCVCTETITKAGRGMPYKYKYVTKGLMPTKTYITLMDIYPADDRKRYKFAENKMS	122			
Db	1	LKTSDDITETITETITKAGRGMPYKYKYVTKGLMPTKTYITLMDIYPADDRKRYKFAENKMS	60			
Qy	123	VTGKAEPAPCRILYVHPDSPATGACNMNOVYSPQKLTLTNNHLDPRCHILNSMXYKQPR	182			
Db	61	VTGKAEPAPCRILYVHPDSPATGACNMNOVYSPQKLTLTNNHLDPRCHILNSMXYKQPR	120			
Qy	183	LHLYVKAEDNNGFSCNKAFCCTVHPPEFAIVYVSYONKKTOLKLTENNPFAGKR	237			
Db	121	LHLYVKAEDNNGFSCNKAFCCTVHPPEFAIVYVSYONKKTOLKLTENNPFAGKR	173			
RESULT 14						
AC	091B66	PRELIMINARY	PRT:	132 AA.		
DT	01-OCT-2000	(TRENDELrel. 15, Created)				
DT	01-OCT-2000	(TRENDELrel. 15, Last sequence update)				

```

DT 01-MAR-2002 (TREMblrel. 20. Last annotation update)
DE Transcription factor Tbx4 (Fragment).
GN Tbx4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20171077; PubMed=10704879;
RA Takabatake Y., Takabatake T., Takeshima K.;
RT "Conserved and divergent expression of T-box genes Tbx2-Tbx5 in
RT Xenopus.";
RL Mech. Dev. 91:433-437(2000).
DR EMBL: AB032943; BAA93083.1; -.
DR HSSP: P24781; 1XBR.
DR TRANSFAC: T04421; -.
DR InterPro: IPR001699; TF_T-box.
DR Pfam: PF00907; T-box.1.
DR PRINTS: PR00937; TBOX.
DR SMART: SM00425; TBOX.1.
DR PROSITE: PS01264; TBOX_2; 1.
DR PROSITE: PS0252; TBOX_3; 1.
FT NON_TER 1
FT TER 132
SQ SEQUENCE 132 AA; 14978 MW; E8C79EC8B83466 CRC64;

Query Match 34.8%; Score 656.5; DB 13; Length 132;
Best Local Similarity 91.0%; Pred. No. 4,5e-52;
Matches 121; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

OY 118 DNKMSYTCGAEPAMGRLVYHDPSPATGAMHMQLVSPQKLTLNHMLDPFGHIIINSMH 177
DB 1 DNKMWVAKKEPAMGRLVYHDPSPATGAMHMQLVSPQKLTLNHMLDPFGHIIINSMH 60
OY 178 KYOPRLHIVAKDENNGSGSNAPACTCTTVPETAFIANTSYONHKITOLKIENNPFAKGR 237
DB 61 KYOPRLHIVAKDENNAFCSNAPACTCTTVPETAFIANTSYONHKITOLKIENNPFAKGR 120
OY 238 GSDMELHRRMSM 250
DB 121 GSDSDL-RVARL 132

RESULT 15
ID ORUW76 PRELIMINARY; PRT: 440 AA.
AC ORUW76;
DT 01-MAR-2002 (TREMblrel. 20. Created)
DT 01-MAR-2002 (TREMblrel. 20. Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21. Last annotation update)
DE Transcription factor Tbx20.
GN Tbx20.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21850180; PubMed=11862462;
RA Iio A., Koide M., Hida K., Morisaki T.;
RT "Expression pattern of novel chick T-box gene, Tbx20.";
RL Dev. Genes Evol. 211:559-562(2001).
DR EMBL: AB070554; BAB83622.1; -.
DR InterPro: IPR001699; TF_T-box.
DR Pfam: PF00907; T-box.1.
DR PRINTS: PR00937; TBOX.
DR SMART: SM00425; TBOX.1.
DR PROSITE: PS01283; TBOX_1; UNKNOWN_1.
DR PROSITE: PS0252; TBOX_3; 1.
SQ SEQUENCE 440 AA; 48687 MW; 3F8AF3E327C0930C CRC64;

```

```

Query Match 34.6%; Score 654; DB 13; Length 440;
Best Local Similarity 43.8%; Pred. No. 4e-51;
Matches 149; Conservative 48; Mismatches 99; Indels 44; Gaps 10;

OY 14 PLEPPAKDLCDCKSPESAL-----GAPSKSPSPQAAFT----OQCEGIVFLH 59
DB 41 PLEQFEKSSC-AOPLSDLGLPEHGFSGSPSALCTEDLIPPTPIIPSEEMAKISCISLE 99
OY 60 ERELMLKFEHVCTEMITTKAGRMFPSPYKVTGLNFKTKYTLIMDTVPADDRKVFADN 119
DB 100 TKELMDKFHELCTEMITTKSGRMFPPIRSPSGVDEAKYIVLMDIYVDNKKRYAYH 159
OY 120 K--WSYTGKAEPMGRLVYHDPSPATGAMHMQLVSPQKLTLNHMLDPFGHIIINSMH 177
DB 160 RSSMLVAKKADPELRLVYHDPSPATGALQKQVSPQKLTLNHMLDPFGHIIINSMH 219
OY 178 KYOPRLHIVAKDENNG--FGSKNTAFCTHVPEETAFAVTSYONHKITOLKIENNPFAK 235
DB 220 KYOPRLHIVAKDENNG--FGSKNTAFCTHVPEETAFAVTSYONHKITOLKIENNPFAK 279
OY 236 FRGSD--DNELHRR--SRMSKEYPVVPRSTVR----QKVASHSFSSSRALSTSSN 286
DB 280 FRGSRDLIERESVESLQKRSY--ANSPIRYGGEDLDLDDSDATQSGKSAFTTSDN 356
OY 287 L-----GSQYOCENGVSGPSODLLPPNPVPL 313
DB 337 LSLSSWSSSTSPSPGPHQPSLSALGTSTASTATPIPHPI 376

Search completed: June 17, 2003, 09:48:16
Job time : 35 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      June 17, 2003, 09:44:00 ; Search time 39 Seconds
              (without alignments)
              1192.422 Million cell updates/sec
```

Title: US-10-027-059A-1  
 Perfect score: 1889  
 Sequence: 1 MADADEGFGIAHTPLEPDAK.....CDHPWSICFLSYLFLSLGWG 349

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 08
                  Maximum Match 1008
```

Database :  
1. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1980.DAT.\*  
2. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1981.DAT.\*  
3. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1982.DAT.\*  
4. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1983.DAT.\*  
5. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1984.DAT.\*  
6. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1985.DAT.\*  
7. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1986.DAT.\*  
8. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1987.DAT.\*  
9. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1988.DAT.\*  
10. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1989.DAT.\*  
11. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1990.DAT.\*  
12. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1991.DAT.\*  
13. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1992.DAT.\*  
14. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1993.DAT.\*  
15. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1994.DAT.\*  
16. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1995.DAT.\*  
17. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1996.DAT.\*  
18. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1997.DAT.\*  
19. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1998.DAT.\*  
20. /S1S2/gcgdata/geneseq/genescqp-emb1/AA1999.DAT.\*  
21. /S1S2/gcgdata/geneseq/genescqp-emb1/AA2000.DAT.\*  
22. /S1S2/gcgdata/geneseq/genescqp-emb1/AA2001.DAT.\*  
23. /S1S2/gcgdata/geneseq/genescqp-emb1/AA2002.DAT.\*

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	647.5	34.3	716	22	AAW79757	Human protein SPO
2	643.5	34.1	722	22	AAW78773	Human protein SPO
3	626.6	33.1	988	22	ABR60007	Drosophila melanog
4	624	33.0	295	22	ABR50080	Human protein seq
5	618	32.7	699	22	ABR65442	Drosophila melanog
6	595	31.5	360	22	ABR61481	Drosophila melanog
7	585	31.0	463	22	ABR61481	Drosophila melanog
8	582	29.8	424	22	ABR61350	Drosophila melanog
9	553.5	29.3	680	22	ABR62374	Drosophila melanog
10	541.5	28.7	580	22	ABR62371	Drosophila melanog

14	45	95.5	5.1	85	21	AA3C8840
15	44	95.5	5.1	85	21	AA3C8840
16	43	96.5	5.2	971	22	AB6A5836
17	42	98.5	5.2	509	23	AB6A5836
18	41	98.5	5.2	422	24	AA679399
19	39	98.5	5.2	422	24	AA679399
20	38	99	5.2	24	20	AA102871
21	37	101	5.4	872	23	AB676193
22	36	102.5	5.4	872	23	AB676193
23	35	111	5.9	276	32	AB04712
24	34	129	6.8	90	32	AB04712
25	33	158	8.4	37	22	AB036032
26	32	158	8.4	37	22	AB036032
27	31	158	8.4	37	22	AB036032
28	30	158	8.4	37	22	AB036032
29	29	158	8.4	37	22	AB036032
30	28	158	8.4	37	22	AB036032
31	27	158	8.4	37	22	AB036032
32	26	158	8.4	37	22	AB036032
33	25	158	8.4	37	22	AB036032
34	24	177.5	9.5	938	22	AA694129
35	23	179	9.5	156	23	AB651291
36	22	179	9.5	156	23	AB651291
37	21	19.5	10.3	209	22	AB061340
38	20	207.5	10.5	20	20	AA033776
39	19	489	28.9	747	22	AB6B3134
40	18	499.5	26.4	517	21	AA574414
41	17	505	26.7	688	21	AA574414
42	16	507	26.8	688	21	AA574414
43	15	513	27.2	448	20	AA573793
44	14	517.5	27.4	365	21	AA573793
45	13	528.5	28.0	535	22	AA651098
46	12	535	28.3	423	23	AA651098
47	11	535.5	28.3	423	23	AA651098
48	10	535	28.3	423	23	AA651098
49	9	535	28.3	423	23	AA651098
50	8	535	28.3	423	23	AA651098
51	7	535	28.3	423	23	AA651098
52	6	535	28.3	423	23	AA651098
53	5	535	28.3	423	23	AA651098
54	4	535	28.3	423	23	AA651098
55	3	535	28.3	423	23	AA651098
56	2	535	28.3	423	23	AA651098
57	1	535	28.3	423	23	AA651098
58	0	535	28.3	423	23	AA651098
59	-1	535	28.3	423	23	AA651098
60	-2	535	28.3	423	23	AA651098
61	-3	535	28.3	423	23	AA651098
62	-4	535	28.3	423	23	AA651098
63	-5	535	28.3	423	23	AA651098
64	-6	535	28.3	423	23	AA651098
65	-7	535	28.3	423	23	AA651098
66	-8	535	28.3	423	23	AA651098
67	-9	535	28.3	423	23	AA651098
68	-10	535	28.3	423	23	AA651098
69	-11	535	28.3	423	23	AA651098
70	-12	535	28.3	423	23	AA651098
71	-13	535	28.3	423	23	AA651098
72	-14	535	28.3	423	23	AA651098
73	-15	535	28.3	423	23	AA651098
74	-16	535	28.3	423	23	AA651098
75	-17	535	28.3	423	23	AA651098
76	-18	535	28.3	423	23	AA651098
77	-19	535	28.3	423	23	AA651098
78	-20	535	28.3	423	23	AA651098
79	-21	535	28.3	423	23	AA651098

## ALIGNMENTS

RESULT 1  
ADM79757

ID AAM79757 standard; Protein; 716 AA.

AC AAM79757

DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 3403

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.

05 Homo sapiens

PN W0200157190-A2

PD 09-AUG-2001.

05-FEB-2001; 2001WO-US04098

PR	03-FEB-2000;	200005-0456914
PR	27-FEB-2000;	200005-0560875
PR	21-APR-2000;	200005-0558075
PR	20-JUN-2000;	200005-0620325
PR	19-JUL-2000;	200005-0654936
PR	01-SEP-2000;	200005-0663361
PR	15-OCT-2000;	200005-0693325
PR	28-SEP-2000;	200005-0728422
PR	30-NOV-2000;	200005-0728422
XX		
PR	(HYSE-) HISEO INC.	









CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 360 AA:

Query Match 31.5%; Score 595; DB 22; Length 360;

Best Local Similarity 45.3%; Pred. No. 5,le-50;

Matches 124; Conservative 42; Mismatches 92; Indels 16; Gaps 5;

39 PSSPQAAFTQOGMEGIVFLHERLMLKPHFVGEVMTITKAGRRMPSPYKVKYKLTGILNPKT 98  
DB 20 PSRPT-----LPGVEAKLNNNDLMQOFHRTIGETMITTKSGRRMPSPMRVSLGLEFEA 72  
OY 99 KYILLMDIYPADDDHRYFAADNKNASVTKAEAPMGRGLVHPDSDPATGAHMMROLVSPQK 158  
DB 73 SYCVLELMPVYIDDCRYKFGSQWVPAGCAEPQSPQRMVLPDSDPATGAHMQSOLLFPKY 132  
OY 159 KLTNNHLDPRGHIIILNSMKRYOPRLHIYKADENNGF-GSKNTAFCTHVPEFTAFVATVSY 217  
DB 133 KLTNNLTDSGGIVYLAEMKRYOPRLHITSELTOLPMAPOAF--VFPELTFVAVYAA 189  
OY 218 OHNKTQTLKIENPPAKFGSDDMELHRMSQSKYPPVPRSTVROKVASNHSPPSE 277  
DB 190 QNDRTITLAKIDNPPAKFGRESGOSRCK-KLNSSGNSLTSESDGSSVSCDSPQAKR 247  
OY 278 STALSTSNLSQYOCENGVSQSDLLPPEPY 311  
DB 248 OHDPFEDSDTGSVSPATVGAHP---VANPMSPT 278

RESULT 7

ID ABB61330 standard; Protein; 469 AA.

XX ABB61330;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 10782.

XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EM;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL05433.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -

XX Disclosure; SEQ ID NO 10782; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention is

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (AB57737-AB872072).

CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 469 AA:

Query Match 31.0%; Score 595; DB 22; Length 469;

Best Local Similarity 43.5%; Pred. No. 7,4e-49;

Matches 130; Conservative 44; Mismatches 81; Indels 44; Gaps 8;

OY 51 MEGIVFLHERLMLKPHFVGEVMTITKAGRRMPSPYKVKYKLTGILNPKTITLMDIYPAD 110  
DB 53 LGCEVMTIQLDNDLMQOFHRTIGETMITTKSGRRMPSPMRVSLGLEFSNVCILLEMPYD 112  
OY 111 DHRVFAADNKNASVTKAEAPMGRGLVHPDSDPATGAHMMROLVSPQKLTNNHLDPRGH 170  
DB 113 DCRVFGSQWVPAGCAEPQSPQRMVLPDSDPATGAHMQAOPLEFKYKLTNNLTDSNGH 172  
OY 171 IILNSMKRYOPRLHIYKADENNGF-GSKNTAFCTHVPEFTAFVATVSYONHKTQTLKEN 229  
DB 173 IYLAEMKRYOPRLHIYRTADLAQIPMAPOAF--VFAETEFVAVYAYOYDRITLKTIDN 229  
OY 230 NPPAFGR-----GSDMDLHRMSQSK-EYVPPV-----RST 262  
DB 230 NPPAFGRFRETGOSRCKRRKSSPTGSDQSPSLSHSQSDESDMSPTKGCGCTAGCTSS 289  
OY 263 VROKVASNHSPPSESRALSTSNLSQYOCENGVSQSDLLPPEPY 313  
DB 290 IGDSGDPQIKRLRNSGACSLSSSDQ-----SYGASLALGASPPPHLHSHHPHL 343

RESULT 8

ID ABB61258 standard; Protein; 424 AA.

XX ABB61258;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 10566.

XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EM;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL05361.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -

XX Disclosure; SEQ ID NO 10566; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention sequences (AM018440-AM16175) and the encoded proteins (AB55777-AB572022). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPAC at <http://wipo.int/pub/published/PCL/sequences>.

**SQ Sequence 424 AA;**

Query Match	29.88;	Score 562;	DB 22;	Length 424;
Best Local Similarity	37.88;	Prod No. 13046;		

Matches 133; Conservative 41; Indels 84; Gaps 10

[illegible]

RESULT 9  
ABB6252A

AC ABB62524;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster SEQ ID NO 14364.

Drosophila; developmental biology; cell signaling; insecticide

XX  
OS  
*Drosophila melanogaster*.

XX WO200171042-A2.  
PN

XX  
PD 27-SEP-2001.

AA  
PF 23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

XX 11 JUL 2000, 200005-00441000

(FIVE) FE CONF IN:

XX

XX

XX	VERICEY, BO	VERICEY, BO	VERICEY, BO
XX	VERICEY, BO	VERICEY, BO	VERICEY, BO

DR N-PSDB; ABL066627.

PT	New Isolated nucleic
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

PT genes from *Drosophila* and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure: SEQ ID NO 14364; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent.

[illegible]

**SQ** Sequence 580 AA;

Query Match	29.3%	Score 553.5;	DB 22;	Length 580;
Best Local Similarity	37.4%	Pred. NO. 1	3e-45;	

Matches 125; Conservative 54; Mismatches 96; Indels 59; Gaps 10;

QY	24	CDSR	PESALAGAKSKSSDD-----AAFPQMGAGKVCYLADEHEMLAKHEVGHEN	704
QY	147	GST	POSOTGTDEBERLITPPPOAKAPRYKESGCMODLLAPVOCHLETTELMOANHIDIGTEN	705
QY	75	IIITAGHMRHPSFKXXVYTG-----LNPRTYITLLMDIYADODHRRKFNADK-----NISTYGAIA	128	
Db	207	IIITGSRHPTATRSFSGCLYLOIADPDARYVLMIIIDMSRRRYAYRSAPLAKAD	266	
QY	129	PAMQRLVHPDSPTATGAMHMRQVLSFORLKTNRNIDPFCHILNLSMKHYQDRLHIVYA	188	
Db	267	PAPRALVIMHDSFSGCALAQVISTEVRVYLNLSMKKQDVIYLSMKHYQRIHILHIV	326	
QY	189	DENSGPRS-----KNTACHPHFPEPTAFVATYSQMHITQULKEINPPACGPRS-----	239	
QY	327	SHGOSTPSPKEIJDLDHMKTYFFPEVPAATYANOGLITKLKIDSPPACGROSRLT	386	
Db	240	-----DDME-----LHRMSKQSPVYKRSIVYRKVAYSNISPPSSSESLSTSSN-----	286	
Db	387	DPVMDPHEMLLLDQMSPLATLTF-----DVIJLAKAGDPP-----SSMLPFXKARQIHQI	439	
QY	287	-----LGSQYQCEANQVSGSDLLPPPP	310	
QY	440	GQNSPVALHMPQROAAAALAGP-----pppp	468	

RESULT 10 .

ID ABB62471

AC ABB62471;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 14205.

**Drosophila; developmental**

2000

XX  
DN 10300171043-XX  
PD 27-SEP-2001

XX 22-MAR-2001: 2001WO-11509221  
DE

XX  
23-MAR-2000: 2000RS-191637P

11-JUL-2000; 2000US

Pt	XI	Venter JC, Adams M, Li FWD, Myers EW;
XX		
DR	XI	WPI; 2001-656860/75.
XX		
DR	N-Psdb; ABL06574.	
XX		
Pt	Gen isolated nucleic acid detection reagent for detecting 1000 or more	
Pt	genes from Drosophila and for elucidating cell signalling and cell-cell	
PT	interactions -	
XX		
PS	Disclosure; SEQ ID NO 14205; 21bp + Sequence Listing; English.	
XX		
CC	The invention relates to an isolated nucleic acid detection reagent	
CC	capable of detecting 1000 or more genes from Drosophila. The invention is,	
CC	useful in developmental biology and in elucidating cell signaling and	
CC	cell-cell interactions in higher eukaryotes for the development of	
CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
CC	discloses genomic DNA sequences (AB116176-AB120511), expressed DNA	
CC	sequences (AB101840-AB116175) and the encoded proteins	
CC	(AB157731-AB170212).	
CC	The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WFO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 660 AA;	
	Query Match 28.7%; Score 541.5; DB 22; Length 660;	
	Best Local Similarity 37.7%; Pred. No. 2.5e+44;	
	Matches 126; Conservative 57; Mismatches 120; Indels 31; Gaps 10;	
OY	9 GLAHPLEPDADKLPDSKPESALGAPSKSPS---SFQ--AAFTQGMEGRVFLHEREL 63	
D*	229 GAATPPEPRPPR--PSOSPEERLSPSSPAQDPRTIVGSNCNCDLTTPVOCHLETREL 286	
OY	64 WLKEHENGEMTHITTAARRRRPFSTKYKVYG---LNREKTILLMDIVADHRKRPPADN 119	
D*	287 WKRFHELGATMIIITSGRRNFPPAVSFSGLROIPADRYALLDVPDLSRRIRAYH 346	
OY	120 K--NSYTGAKAEFRAMRGILYYHSDPATGAGAHMKROLVSFOKLKATINHLDPGHITLNSKH 177	
D*	347 RSMVLNAGKKDDPPRSKITVAHDPCSLSEALKQGVSEVKYLTKNNKDSGGVLYANSHK 406	
OY	178 KYORPHIUYAADENNGF--GS---KNARFCITYPETRAFIATYSTONHKITLOLKTIENNPF 232	
D*	407 KQRPRIHYLVLSHGOSTFGSPKELDNDHTTVEFYETFAVAYAQMLITKLIKIDSNPF 466	
OY	233 AGCGFSGDDDELHRASMSOSES---KEPVVPSTYVRCKVANSKHSFPSSSRALSITSSNLG 288	
D*	467 AAGCFDSSRLSDGFDDDMDAFFFDDBHRTKPLAFRRPDLMSOLUTQDADAASHALLXK 526	
OY	289 SOYGCENVSGRPSODLLP-----PPNYXP 314	
D*	527 QHLQK-FGRSPTYEKLPLHYGRSAAPRRPPRAP 559	
XX		
RESULT 11		
ID	AAA84794 standard; Protein: 423 AA.	
AC	AAA84794;	
XX		
DT	22-JUL-2002 (first entry)	
XX		
DE	zebrafish ntl.	
XX		
KW	Promoter; ribozyme; RNA polymerase; zebrafish; morphological;	
KW	behavioral; change; light tropism; chemical tropism; identical;	
KW	disease related gene; ntl; mouse; brachyury; xbra; Xenopus.	
OS	Brachydanio rerio.	
XX		
US	US6355415-B1.	

[illegible]

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rag

Page 8

```
AC AAB51014.
XX 20-MAR-2001 (first entry)
XX
XX Murine T-bet protein.
XX
XX Murine: T-bet: T box expressed in T cells protein; autoimmune disease;
XX Th2-type dysfunction.
XX
XX Mus musculus.
XX
XX M0200073453-A1.
XX
XX 07-DEC-2000.
XX
XX 01-JUN-2000: 2000MO-US15345.
XX
XX 02-JUN-1999: 990US-0137085.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Glimcher LH, Szabo SJ:
XX
XX NPI: 2001-061537/07.
XX
XX N-PSDB: AAC91989.
XX
XX Novel protein, T-bet, a new member of T-box family of transcription
XX factors is useful for identifying modulators of T-bet activity which
XX treat allergies, cancers, autoimmune and infectious diseases -
XX
XX Claim 17: Page 123-124; 126pp: English.
XX
XX
XX The present sequence is murine T box expressed in T cells (T-bet)
XX protein. T-bet can be used in the treatment of autoimmune diseases
XX associated with the T-bet protein, such as rheumatoid arthritis,
XX associated alopecia, autoimmune thyroiditis, psoriasis,
XX keratoconjunctivitis, alopecia areata, allergic responses due to
XX arthropod bite reactions, Crohn's disease, aphthous ulcer, lillitis,
XX conjunctivitis, keratoconjunctivitis, ulcerative colitis, asthma,
XX allergic asthma, cutaneous lupus, erythematous, scleroderma, vaginitis,
XX differentiation of th cells.
XX
XX Sequence 530 AA:
XX
XX Query Match 28.3%; Score 535; DB 22; Length 530;
XX Similarity 35.2%; Pred. No. 35; Indels 134; Gaps 16;
XX Matches 147; Conservative 40; Mismatches 97;
XX
XX 16 EPDADIDPCDSKPSALGAP-----SKSPSSQAAAF-----T00 49
XX
XX 38 EPDADIDPCDSKPSALGAP-----SKSPSSQAAAF-----T00 49
XX
XX 50 CMGC-----IVFLMERELMLKHFENGTEN 74
XX
XX 97 GAGCGPVVDCYAPADPRAGILPGREDYALPAGLEVSKLVALNHLILMSKFNQHOTEM 156
XX
XX 75 IITKAGRNPSYKVVYGLAPKTKYLLMDIVPADDRYKFAUNKSVYGAERPAAG- 133
XX
XX 157 IITKAGRNPSYKVVYGLAPKTKYLLMDIVPADDRYKFAUNKSVYGAERPAAG- 216
XX
XX 134 RLYVHPSPSPATGAMHNRQLVSFQKLTNTN---HLDPFGHLLNSMNYOPRLHIVADNK 190
XX
XX 217 RLYVHPSPSPATGAMHNRQLVSFQKLTNTN---HLDPFGHLLNSMNYOPRLHIVADNK 276
XX
XX 191 NMGFGSKNAPC---THV--PEPFAFIATVYONKHITOLKLENNPAAGR----- 237
XX
XX 277 ---GPEFACASNTYHVFDEOFIAVTVYONKHITOLKLENNPAAGR----- 331
XX
XX 238 -----GSDMLKNSNOSKERV-----VBSGTROK--- 266
XX
XX 332 ASVDTSPSPPCNCLGDDP-----FSLNLSQPVHSPHFYRPLQPKIMLSQPYVL 386
```

```
OY 267 -VASNHPSPSPESRALSTSSNIGSOVCENGVSGPS-----CDLLPPNRYDLPOEH 317
DB 387 GTPREHS-YEAERFRAVSMKPTL-----LPASCGPTVYRYRGQDVLAPQAGMVPAPOT 437
XX
XX ResurT 13
XX ID AAB51013 standard; protein: 535 AA.
XX
XX AAB51013:
XX
XX 20-MAR-2001 (first entry)
XX
XX Human T-bet protein.
XX
XX Human: T-bet: T box expressed in T cells protein; autoimmune disease;
XX Th2-type dysfunction.
XX
XX Homo sapiens.
XX
XX M0200073453-A1.
XX
XX 07-DEC-2000.
XX
XX 01-JUN-2000: 2000MO-US15345.
XX
XX 02-JUN-1999: 990US-0137085.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Glimcher LH, Szabo SJ:
XX
XX NPI: 2001-061537/07.
XX
XX N-PSDB: AAC91989.
XX
XX Novel protein, T-bet, a new member of T-box family of transcription
XX factors is useful for identifying modulators of T-bet activity which
XX treat allergies, cancers, autoimmune and infectious diseases -
XX
XX Claim 16: Page 118-120; 126pp: English.
XX
XX
XX The present sequence is human T box expressed in T cells (T-bet) protein.
XX T-bet can be used in the treatment of autoimmune diseases associated with
XX a Th2-type dysfunction e.g. diabetes mellitus, keratoconjunctivitis,
XX associated alopecia, autoimmune thyroiditis, psoriasis,
XX alopecia areata, allergic responses due to arthropod bite reactions,
XX Crohn's disease, aphthous ulcer, lillitis, conjunctivitis,
XX keratoconjunctivitis, ulcerative colitis, asthma, allergic asthma,
XX cutaneous lupus, erythematous, scleroderma, vaginitis, and
XX drug eruptions. The role of T-bet is to drive differentiation of th
XX cells.
XX
XX Sequence 535 AA:
XX
XX Query Match 28.0%; Score 528.5; DB 22; Length 535;
XX Similarity 37.7%; Pred. No. 35; Indels 115; Gaps 12;
XX Matches 140; Conservative 41; Mismatches 97;
XX
XX 3 DABE--GGIAATPLPEADKIDPCDSKPSALGAPSKSPSSQAAAF-----HE 52
XX
XX 44 DABEIRGGSLSQSY--PGSALVP--APPSHFATVAVPRPQAGPQAGSEFFPPADAE 100
XX
XX 53 C-----IKYFLMERELMLKHFENGTENIT 77
XX
XX 101 GYQPEQCYAAPPRAGILPGREDYALPAGLEVSKLVALNHLILMSKFNQHOTEMIT 160
XX
XX 78 KAGRNPSPSYKVVYGLAPKTKYLLMDIVPADDRYKFAUNKSVYGAERPAAG-RLY 116
XX
XX 161 KAGRNPSPSYKVVYGLAPKTKYLLMDIVPADDRYKFAUNKSVYGAERPAAG-RLY 220
XX
XX 137 VHPDSPATGAMHNRQLVSFQKLTNTN---HLDPFGHLLNSMNYOPRLHIVADNKNG 193
XX
XX 221 VHPDSPATGAMHNRQLVSFQKLTNTN---HLDPFGHLLNSMNYOPRLHIVADNKGP 280
```

OY 194 FGSKNATFATYHV--PPEAFIAVTSYONKRIKTIENNPFAKGRFSODMELHRMSMO 251  
 DB 281 EAAHC-ASNNHIFTEFOETOFIAVAYONAEITOLKIDNNPFAKCFREN----PESWYTSV 335  
 OY 232 SKKEYVPRSTYRKVKVANSNPSFSESRLSTSSNLSGQYCCER---GVSGSPDILPP 307  
 DB 336 DTSPSPGNCQFLGSDHYSPD-----LPNQVPYPRRYPOLPGQAKDVVQ 383  
 OY 308 PNPYPLDQHS 318  
 DB 384 AYWLGAPEHDS 394

RESULT 14  
 AAB57098  
 ID AAB57098 standard; Protein: 365 AA.  
 XX AAB57098;  
 AC  
 DT 13-MAR-2001 (first entry)  
 XX  
 DE Human prostate cancer antigen protein sequence SEQ ID NO:1676.  
 XX  
 KW Human: prostate cancer; prostate cancer antigen; detection; diagnosis;  
 KW neuroprotective; cytostatic; cardiostatic; immunomodulatory; muscular;  
 KW vulnereary; gastrointestinal; nephrotoxic; antineoplastic; gynaecological;  
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
 KW gastrolntestinal; pulmonary; cardiovascular; proliferative disorder;  
 KW wound; infectious disease.  
 XX  
 OS Homo sapiens.  
 XX  
 FN W020005174-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 \* 08-MAR-2000; 2000MO-USO5988.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 XX (HMAN-) HUMAN GENE SEI INC.  
 PA (ROSE/) ROSEN C A.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI: 2000-587513/55.  
 XX  
 N-PSDB: AAF16301.  
 XX  
 PT Prostate cancer associated gene sequences, referred to as prostate  
 PT cancer antigens, useful for treatment, prevention, and diagnosis of  
 PT disorders such as prostate cancer -  
 XX  
 PS Claim 11: Page 2144-2145; 2338pp: English.  
 XX  
 CC AAF15566 to AAF16505 encode the human prostate cancer associated  
 CC proteins, called prostate cancer antigens, given in AAB56563 to AAB57302.  
 CC The prostate cancer antigens can have neuroprotective, cytostatic,  
 CC cardiostatic, immunomodulatory, muscular, vulnereary, gastrointestinal,  
 CC nephrotoxic, antineoplastic, gynaecological and antibacterial activities,  
 CC and can be used in gene therapy. The prostate cancer antigen  
 CC polynucleotides may be used for detection of prostate cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The prostate cancer antigens may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
 CC AAB57303 represent sequences used in the exemplification of the present  
 CC invention.  
 CC  
 SO Sequence 365 AA;  
 Query Match 27.4%; Score 517.5; DB 21; Length 365;

Best Local Similarity 45.9%; Pred. No. 2,4e-42;  
 Matches 112; Conservative 33; Mismatches 74; Indels 25; Gaps 4;  
 OY 83 MPFSYKVVYVGNLPMPTXYILLMDIVPADHRRKFAADKKSVYTGKAPAMPGRILYHPDSP 142  
 DB 4 MPEPKVRCGSLDCKKAXYILLMDIADDCRYKRNHMYAGADPEMPRYTHPSP 63  
 OY 143 ATCAIHMMOLVPSFOKLKLTINILDPQC-----HIIINSHKYOPR 182  
 DB 64 ATQEQMSKVVFPFKLKTINISDKHGFTLAFPSDATHQGCNVSFCGTQILINSHKYOPR 123  
 OY 183 LHTVADENNGFSGSKNPAFCCTVPPPTAFVATSYONKRIKTIENNPFAKGRFSODM 242  
 DB 124 FHTVRA--NDILKLYSPFTYLPETEFIAVTAQDKITOLKIDNNPFAKGRFDGNG 181  
 OY 243 ELHRMSRMOSKEYPVPRSTYRKVKVANSNPSFSESRLSTSSNLSGQYCCENGVSQSQ 302  
 DB 182 RREKROLTLQSMRVFDERHKKEKNTSDSE--SSEQAFNFAQASSPASTVGTSEN-LK 238  
 OY 303 DLLP 306  
 DB 239 DLCP 242

RESULT 15  
 AAY03773  
 ID AAY03773 standard; Protein: 448 AA.  
 XX AAY03773;  
 AC  
 DT 11-JUN-1999 (first entry)  
 XX  
 DE Human T-Box polypeptide, Tbx19.  
 XX  
 KW T-Box polypeptide; Tbx19; Holt-Oram syndrome; Uinar Mammary syndrome;  
 KW Digeorge syndrome; cancer; psychiatric disorder; heart defect; human.  
 XX  
 OS Homo sapiens.  
 XX  
 FN W09911783-A2.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PR 02-SEP-1998; 98WO-EP05713.  
 XX  
 PR 17-AUG-1998; 98GB-0017900.  
 PR 03-SEP-1997; 97EP-0306830.  
 XX  
 PA (SMK ) SMITHKLINE BEECHAM PLC.  
 PA (GONO-) UNIV NOTTINGHAM.  
 XX  
 PI Brook JD, Terrett JA, Yi CH;  
 XX  
 DR WPI: 1999-205187/17.  
 XX  
 N-PSDB: AAX31841.  
 XX  
 PT Tbx19 protein - used to treat Holt-Oram syndrome, Uinar Mammary  
 PT syndrome, Digeorge syndrome, cancer, psychiatric disorders, and  
 PT heart defects  
 XX  
 PS Claim 3: Page 23; 32pp: English.  
 XX  
 CC This represents a T-Box polypeptide having Tbx19 activity. Host cells  
 CC containing a vector comprising the Tbx19 nucleic acids are used for the  
 CC recombinant expression of the protein. The Tbx19 polypeptide can be used  
 CC for treating Holt-Oram syndrome, Uinar Mammary syndrome, Digeorge  
 CC syndrome, cancer, psychiatric disorders, and heart defects.  
 CC  
 SO Sequence 448 AA;  
 Query Match 27.2%; Score 513; DB 20; Length 448;  
 Best Local Similarity 39.2%; Pred. No. 9,3e-42;  
 Matches 113; Conservative 43; Mismatches 100; Indels 32; Gaps 5;

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rag

Page 10

OY 54 IVEVHERELMLKFEVGTETITKAGRRPSPYKVYVGLAPKTKYLLMDLVPADDR 113  
DB 41 IQLIEDAPLMORKEVITNEMLYTKNGRMEVYAKISYDOPNMYSLDVPFTDSHR 100  
OY 114 YFPAOKMSYTGKAEPMGRLYVHPDSPATGAHMQLVSEFOKLJTNHLDPCGHITL 173  
DB 101 KRYVNGEMVPAGKPEVESHCVYIHPDSPFGAHMKAPISFSKVKLTKR-LNGGQIML 159  
OY 174 NSMKTQPLHIYKADENNGFSKNTAFCTIVPEPAFIAYSTONKTIQKTIENPFA 233  
DB 160 NSMKTQPLHIYKADENNGFSKNTAFCTIVPEPAFIAYSTONKTIQKTIENPFA 213  
OY 234 KGFSGDDDELHMRSMQSKREYVVPSTYRQKVASNHSFSSSRALSTSSNLGSOYOC 293  
DB 214 KAF-----LDAKERNHLDVPEA-----ISEOHVYTSHLGWMIFSN 250  
OY 294 ENGV--SGFSODLLPPNPPLQPHSOTYHCTKRKGCDHMAICPL 339  
DB 251 PDCVCTAGNSHWQYAAPLPLPAPHTHGCHEHYSGLRGHROAPYPSAVM 298

Search completed: June 17, 2003, 09:49:01  
Job time : 41 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:47:23 ; Search time 14 Seconds  
(without alignments)  
733.471 Million cell updates/sec

Title: US-10-027-059A-1  
1889  
Sequence: 1 MADADGEGLAHHTPEPPDAK.....CDHPWSICFLSTLSLGMG 349

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents\_Ah+  
1: /cgn2\_6/ptcdat4/1/1aa/3A.COMB.pep: +  
2: /cgn2\_6/ptcdat4/1/1aa/3B.COMB.pep: +  
3: /cgn2\_6/ptcdat4/1/1aa/3A.COMB.pep: +  
4: /cgn2\_6/ptcdat4/1/1aa/3B.COMB.pep: +  
5: /cgn2\_6/ptcdat4/1/1aa/3A.COMB.pep: +  
6: /cgn2\_6/ptcdat4/1/1aa/3B.COMB.pep: +

- Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	535.5	28.3	423	4	US-08-939-366-6
2	510	27.0	567	3	US-09-188-811-2
3	506.5	26.8	517	3	US-09-189-760-2
4	506.5	26.8	517	3	US-09-514-422-2
5	99	5.2	24	4	US-09-227-357-522
6	98.5	5.2	411	4	US-08-470-339-189
7	98.5	5.2	422	4	US-08-467-602-324
8	98.5	5.2	456	4	US-08-467-602-366
9	98.5	5.2	601	4	US-08-467-602-323
10	98.5	5.2	601	4	US-08-467-602-323
11	98.5	5.2	610	4	US-08-470-335-236
12	98.5	5.2	635	4	US-08-467-602-365
13	98.5	5.2	644	4	US-08-467-602-374
14	98.5	5.2	818	4	US-08-470-335-234
15	98.5	5.2	818	4	US-08-467-602-321
16	98.5	5.2	827	4	US-08-470-335-237
17	98.5	5.2	827	4	US-08-467-602-333
18	98.5	5.2	861	4	US-08-467-602-363
19	98.5	5.2	861	4	US-08-470-335-235
20	98.5	5.2	865	4	US-08-467-602-322
21	98.5	5.2	865	4	US-08-470-335-238
22	98.5	5.2	874	4	US-08-467-602-334
23	98.5	5.2	889	4	US-08-467-602-364
24	98.5	5.2	908	4	US-08-467-602-376
25	94.5	5.0	422	1	US-08-036-555B-170

28	94.5	5.0	422	1	US-08-469-569-170	Sequence 170, App
29	94.5	5.0	422	1	US-08-488-926-3	Sequence 3, App1
30	94.5	5.0	422	1	US-08-249-322A-170	Sequence 170, App
31	94.5	5.0	422	1	US-08-488-927-3	Sequence 3, App1
32	94.5	5.0	422	1	US-08-488-298-3	Sequence 3, App1
33	94.5	5.0	422	1	US-08-339-517-3	Sequence 3, App1
34	94.5	5.0	422	1	US-08-469-526A-170	Sequence 170, App
35	94.5	5.0	422	2	US-08-734-644A-170	Sequence 170, App
36	94.5	5.0	422	2	US-08-469-660-170	Sequence 170, App
37	94.5	5.0	422	4	US-08-341-018-72	Sequence 72, App1
38	94.5	5.0	422	4	US-08-470-335-170	Sequence 170, App
39	94.5	5.0	422	4	US-08-735-021-170	Sequence 170, App
40	94.5	5.0	422	4	US-08-734-644A-170	Sequence 170, App
41	94.5	5.0	422	4	US-08-470-339-170	Sequence 170, App
42	94.5	5.0	422	4	US-08-467-602-170	Sequence 170, App
43	94.5	5.0	422	5	PCT-US94-05083C-165	Sequence 165, App
44	94.5	5.0	422	5	PCT-US94-05083C-185	Sequence 185, App
45	94.5	5.0	422	5	PCT-US95-06846A-170	Sequence 170, App

## ALIGNMENTS

```

RESULT 1
US-08-939-366-6
: Sequence 6, Application US/08939366
: Patent No. 6355415
: GENERAL INFORMATION:
: APPLICANT: Wagner, Thomas E.
: TITLE OF INVENTION: Compositions and Methods for the use of
: TITLE OF INVENTION: Ribozymes to Determine Gene Function
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESS: Medien & Carroll, LLP
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: ZIP: 94102 United States of America
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/939,366
: FILING DATE: 29-SEP-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Carroll, Peter G.
: REGISTRATION NUMBER: 32,837
: REFERENCE/DOCKET INFORMATION: OHU-02749
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 423 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-939-366-6

Query Match      28.3%, Score 535.5, DB 4: Length 423;
Best local similarity 41.3%, Pred. 3.6e-50;
Matches 124; Conservative 40; Mismatches 93; Indels 43; Gaps 9;

QY      40  SQAQAFQOQME-----GKIVLELRLRLKLFHVGTEIITTAGRRFPSPYVAVT 92
      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      15  SAVSESF-QKSGEKDASERDIKLSLDELMTFKETLNEIIVTGTGRHFPVLRAVST 73
      0Y      93  GLNPRTKTYILLIDIVPADHRKRFADKKMSVYKRAEPAGRLYVHPDSPATGAHMRQL 152

```

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rai

Page 2

Db	74	GLDPNAMI	SVLLDPEVAADNDNNKITYVNGVMPGKREPOSPSCVYIHPDSINFLANMKAP	133
Oy	153	VSRQKLTATVNHNLDPGCH	ILINSMKHQPHLTYIADENKOPSKTAFCTHYIPETAT	212
Db	134	AVTSYVYLSMK	LNGGQPMANSLAKREPRITRYI	168
Oy	197	AVTSYQNRKLTQTLQINFEAPKPGKSDIMELHRRMSQSEYRPVPRVSTQKVKASNS	276	
Db	167	AVTAQNEHETLTKHNPKNAP	LDKERSDHEVP	228
Oy	273	PFSSSSHALSTSMK	SOQCEVNGSVGSDPLLPPPPYLPDHSOIVYCTKRRKGC	333
Db	227	TDNQO	SGSOLDGMLPNSGPGSSSS--PQFTGAP-VHSQSYCERYVSILRN	278

```

RESULT 2
US-09-188-811-2
: Sequence 2, Application US/09188811
: Sequence 2, Application US/09188811
: GENERAL INFORMATION:
: APPLICANT: Rhododend, Mehren
: TITLE OF INVENTION: NOVEL MTRX. PROTEIN AND NUCLEIC ACID MOLECULES AND USRS
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: MN1-046CP
: CURRENT APPLICATION NUMBER: US/09/188, 811
: CURRENT FILING DATE: 1998-11-09
: EARLIER APPLICATION NUMBER: 09/163, 116
: NUMBER OF SEQ. ID NOS.: 8
: NUMBER OF SEQ. ID NOS.: 8
: SOFTWARE: Patentlin Ver. 2.0
: SEQ. ID NO. 2
: LENGTH: 567
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-188-811-2

```

[illegible]

```

1 Patent No. 6031078
2
3 GENERAL INFORMATION:
4
5 APPLICANT: Rhodadoust, Mehran
6
7 TITLE OF INVENTION: NOVEL MTXN PROTEIN AND NUCLEIC ACID MOLECULES AND USES
8
9 TITLE OF INVENTION: THEREFOR
10
11 FILE REFERENCE: NMT-046CP2
12
13 CURRENT APPLICATION NUMBER: US/09/189,760
14
15 PENDING PENDING DATE: 1998-11-0
16
17 EARLIER FILING DATE: 1998-09-26, 63,116
18
19 EARLIER APPLICATION NUMBER: 60/089,467
20
21 EARLIER FILING DATE: 1998-06-16
22
23 EARLIER APPLICATION NUMBER: (PENDING)
24
25 EARLIER FILING DATE: 1998-11-09
26
27 NUMBER OF SEQ ID NOS: 10
28
29 SOFTWARE: PatentIn Ver. 2.0
30
31 SEQ ID NO: 2
32
33 LENGTH: 517
34
35 TYPE: PRT
36
37 ORGANISM: Homo sapiens
38
39 IS-09-189-760-2

```

Query Match	Best local similarity	26.38%	Score	506.5	DB 3	Length	517	
		35.8%	Pred.	No. 7.86-7.7				
Matches	Conservative	48	Mismatches	111	Indels	85	Gaps	14
OY	30	SALGAKRSPSSPOLAATPOOGMEI	-----	KYELHREELMLKTHEVGTENMLITKAG	80			
Db	49	SALNARLEPNA-----AGSGCGLOGLGVSPSSFRANHYILCNRPPLMKLHNRHGTENITKOG	-----		105			
OY	81	REMPFSYKXKGTGAMKPTKTYLLMDIYPADODHRYKADKNKSVKQAEKAMPG-RLYVHP	139					
Db	106	REMEFPLSTNIGTNGTNAVYFVUVDLADPHMRPOGCKADKNMGKMYHP	165					
OY	140	DSRPATGAKHMRQLVSTOKLTLNN-----HIDPFGHITLMSMKHYODRLATYKADENN---	192					
Db	166	ESPSTHSGHMKDELSTGKMLKNNKGNANNNTONTIYDLSLKTQPRILATVEYDEYEDL	225					
OY	193	GGSGNATNACHYFPEPTAFATYSYONKHTOLKLENNPRAKSPGSGSDO-----MELHAM	247					
Db	226	NEPKVOTF-----FFSTFOTIATYALVORTDITOLKIDHPFAKPFNDYDSMTLSNDKL	282					
OY	248	SRKOS-----KEYPVVPESTVNYQVNAS-----NHSD-----FSSESRLATSTSSLSOYOCE	294					
Db	283	TVYSPSPSPASPIVPOG-----RIGVGSFPEEPFVNTLQARVYNEENTYAPOTNNLSPOOSE	340					
OY	295	NGSFGSQDILLPPN-----RPGQSHSYIYH-----	332					
Db	341	EVANPDPORILVTPQOTPKNLKLDLSYESEYTSLLPYGIKSLQTSALGAYDPTF	400					
OY	323	-----TRKAKGCDHPMS	335					
Db	401	PAMAGMGKSGSYOKKAAGLPT	423					

```

RESUME
US/09/514-422-2
Sequence 2, Application US/09514422
Patent No. 6291193
GENERAL INFORMATION:
APPLICANT: Khodadoust, Mehran
TITLE OF INVENTION: NOVEL MTX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE OF INVENTION: THEREFOR
FILE REFERENCE: AMT-046CP2
CURRENT APPLICATION NUMBER: US/09/514,422
PRIOR FILING DATE: 1998-02/26/99/189,760
PRIOR APPLICATION NUMBER: 09/163,116
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/089,467
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: (PENDING)

```

RESULT 3  
US-09-189-760-2  
; Sequence 2, Application US/09189760

PRIOR APPLICATION NUMBER: 60/089,466  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: (PENDING)

```

: PRIOR FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 517
: TYPE: prt
: ORGANISM: Homo sapiens
: OS-09-514-422-2

```

Query Match	26.8%;	Score 506.5;	DB 4;	Length 517;
Best Local Similarity	36.3%;	Pred. No. 7.8e-47;		
Matches 139; Conservative	48;	Mismatches 111;	Indels 85;	Gaps 14

```

OY      30  SALGAAKSSPSSFOALFQOOGMECI-----KVILDERELKLNHEGVEEMITTKAG  80
Db      49  SALRAATLEPP-----AAGSCGGLGSLVPGSGFRHAYVLCNRELMLKKNHOETEMITTKG  105
OY      81  RRMPFSYKAVATYLNPKTKYLLIMDIYPADDERHYKFEADKNKNSVGAKEPAPG-RILYHP  139
Db      106  RRMPFSLSPNGLNLTNHTNNEVEVLADPNNHMFEGGKAWTCTGKADNNQGGKATYHP  159
OY      140  DSPATAHNRKROLVSFOKLKLTNN-----HIDPGLILNSMKTPORLATHYKADEN---  192
Db      166  EPRITSSHCHNEHLESTGKLLKNNKGNANNNTOKYVLOSJHTOPRLATHYEDYEDL  245
OY      193  GEGSKTALACRHEPPTAFIATYSOHNKLTOLKLENNPAGRGSDP-----MEIHPM  247
Db      226  NEPSKIQTF---TFSETOFLNATYATQMDLTOLIDNPAGRGMDYDSMTASNDRL  282
OY      248  SRMOS-----KEYPVAPSPRIVOXYS-----NHSP---PFESENALSTSSLSOQOE  294
Db      283  TPSPSTSPSHQIVPGD---RYGVOSFPPEPPVNTLPQARYNGDERVYPTQTNLSPQSE  340
OY      295  NGVSGPSODLLPAPPN-----PY-----PLPOHSQIYH-----  322
Db      341  EVANPQRIALTEVPQDPTGNKLKLDISSYSESYSTLLPYGIKSLPLOTSHALGYDDPFT  400
OY      323  -----OTRKRGEDCDHPMS  335
Db      401  PMAAGMGRCGRSYORKNMAAGLPT  423

```

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rai

Page 4

```

APPLICANT: HILES, IN
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
FILE REFERENCE: US/055,020/02
CURRENT APPLICATION NUMBER: US/08,470/0.399C
EARLIER FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: 08/036,555
EARLIER FILING DATE: 1993-03-24
EARLIER APPLICATION NUMBER: 07/940,389
EARLIER FILING DATE: 1992-09-03
EARLIER FILING DATE: 1992-06-30/07,138
EARLIER FILING DATE: 1992-06-30
EARLIER APPLICATION NUMBER: 07/863,703
EARLIER FILING DATE: 1992-04-03
EARLIER APPLICATION NUMBER: 91 07566.3 GB
EARLIER FILING DATE: 1999-04-10
NUMBER OF SEQ ID NOS: 226
SOURCE AND/OR REFERENCE: Genbank
SEQ ID NOS: 168
LENGTH: 411
TYPE: PAT
ORGANISM: Homo sapiens
US-08-470-339-189

Query Match
Match 177, Conservative 40; Mismatches 119; Indels 99; Gaps 18
5.2% Score 98.5; DB 4; Length 411;
Similarity 23.0%; Pred. No. 0.038;

25 DISPRS:--ALGAPSK--PSPDPAFQDQMGKXVFLHEHLEMLAKHXG 71
Db DREPPAGCRALGPRPEELLANCTGTPMPAFVAGSGEG--EAPRYLKAVOV- 167
OY
114 DREPPAGCRALGPRPEELLANCTGTPMPAFVAGSGEG--EAPRYLKAVOV- 167
Db
72 TEHITKAGRMPSKYKYVGTGLNFKTKILLMDIPADHRKRFADKMSVTKGAEPM 131
OY
168 -----YHAKAGGL--KQDSILVFL-----GTV-----GHFAP 193
Db
132 P--GRL-----YHNDSPATGMMHMQQLYSFOKLTNLNLDPHGKILINSHMKAP 181
OY
134 PSCGLKEDSRITFEMEDNANSTISAPAFRASEPLE-TGRLLKKEVSNTYKRCALP 252
Db
182 RLHIVADENNGFGSKNATFCTVPEPTAFIATVSTQ-----KKTQLTENPFK 234
OY
235 GFSGSDMEHLHRMSRQSKETPVYVRSYVOKVASHNSPSSRSALSTSLQSYQ- 292
Db
233 RLKEKQGP-SAGKSLVLC-ETSSYSLRFTWKNGLNELKRNKPNQITKCP-- 307
OY
308 --GASLTKIKKALSLADSGEYKVIKSLGNDMSANSTI-VESNATSTT-GTSHLVK 362
Db
293 -----GNG-----VSGRQDLLEPPNY 311
OY
363 CAEKETPCVNGCGCHWKKDLSNPRTLKCKPNEF 397
Db

RESULT 7
US-08-467-602-324
Sequence 324, Application US/08467602C
Genbank: US/08467602C
APPLICANT: SKLAR, Robert
APPLICANT: Gayme, David I.
TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND
TITLE OF INVENTION: DISORDERS
FILE REFERENCE: 04585/028003
CURRENT APPLICATION NUMBER: US/08/467,602C
EARLIER FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: 08/036,555
EARLIER FILING DATE: 1994-03-08
EARLIER APPLICATION NUMBER: 08/059,022
EARLIER FILING DATE: 1993-05-06
NUMBER OF SEQ ID NOS: 420
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 324

```

```

1      : LENGTH: 422
2      : TYPE: PRT
3      : ORGANISM: Homo sapiens
4      : US-08-467-602-524
5
6      : Query Match
7      : Best Local Similarity 2.2%; Score 98.5; DB 4; Length 422;
8      : Best Local Similarity 23.0%; Pred. No. 0.039;
9      : Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18
10
11     : 25 DSKPRS-----ALGAPSKS-----PSSPOLAFTQOGMEGIVKFLIERELMLKTFHEVG 71
12     : Db DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
13
14     : 114 DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
15     : Db DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
16
17     : 72 TELLITKAGRIKMPFSKYKVVYMGINPKRTKYLILMDIIVPADDHRYKFRKADNMKSVTQAKERP 131
18     : Db TELLITKAGRIKMPFSKYKVVYMGINPKRTKYLILMDIIVPADDHRYKFRKADNMKSVTQAKERP 131
19
20     : 168 -----NAVKAAGC-----KKDSILVLVL-----GTV-----CHPAF 193
21     : Db 132 P--GRL-----YVHPDPSATGAMKMRQVLVSQFKLKTNNHLDFFGHILNLMHKYOP 181
22     : Db 134 PSQGLLEDSKRTIEFEMPDANSTISRAAPAFASFPPLF--TGRMLKKEYSVLCKRCLAP 252
23
24     : Oy 182 RLHYIKKADENKGNFQKNTAFECSTHVPETAFIATVSYON-----HKTDLKTENMPFAK 234
25     : Db 253 RLKEMKQOE--SAAGSKVLARC--ETSSSESLRFKFWKGNELNRKNRKPQKILQKPF-- 307
26
27     : Db 235 GFRGSDIMELIHKSRMKSKEYEVPVRSYTRAKVNSHSPSSSESRALSTSLNLSQYQ-- 292
28     : Db 308 -----GSHLRIRKNSLADSDGSECMKCYIKSLDADSNANLTI--VESNATVSTIR--GTSLVLYK 362
29
30     : Oy 293 -----CEGNG-----VSPGDDLLPDPNXY 311
31     : Db 363 CAEKETFCVCNCEGCFYKWDLSNRYLCKCPNEF 397
32
33     : RESULT 8
34     : US-08-467-602-366
35     : Sequence 366, Application US/08467602C
36
37     : GENERAL INFORMATION:
38     : APPLICANT: Sklar, Robert
39     : APPLICANT: Marchionni, Mark
40     : APPLICANT: Gymer, David L FOR TREATING MUSCLE DISEASES AND
41     : TITLE OF INVENTION: DISORDERS
42     : FILE REFERENCE: 04585/028003
43     : CURRENT APPLICATION NUMBER: US/08/467, 602C
44     : CURRENT FILING DATE: 1995-06-06
45     : EARLIER FILING DATE: 1994-03-08
46     : EARLIER APPLICATION NUMBER: 08/059, 022
47     : EARLIER FILING DATE: 1993-05-06
48     : NUMBER OF SEQ ID NOS: 420
49     : SOFTWARE: FASTSEQ for Windows Version 4.0
50
51     : SEQ ID NO 366
52     : LENGTH: 456
53     : TYPE: PRT
54     : ORGANISM: Homo sapiens
55
56     : FEATURES:
57     : LOCATION: (34)
58     : OTHER INFORMATION: Xaa is any amino acid
59
60     : US-08-467-602-366
61
62     : Query Match
63     : Best Local Similarity 5.2%; Score 98.5; DB 4; Length 456;
64     : Best Local Similarity 23.0%; Pred. No. 0.045;
65     : Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18
66
67     : 25 DSKPRS-----ALGAPSKS-----PSSPOLAFTQOGMEGIVKFLIERELMLKTFHEVG 71
68     : Db DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
69
70     : 114 DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
71     : Db DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
72
73     : Oy 72 TELLITKAGRIKMPFSKYKVVYMGINPKRTKYLILMDIIVPADDHRYKFRKADNMKSVTQAKERP 131
74     : Db TELLITKAGRIKMPFSKYKVVYMGINPKRTKYLILMDIIVPADDHRYKFRKADNMKSVTQAKERP 131
75
76     : 168 -----NAVKAAGC-----KKDSILVLVL-----GTV-----CHPAF 193
77     : Db 132 P--GRL-----YVHPDPSATGAMKMRQVLVSQFKLKTNNHLDFFGHILNLMHKYOP 181
78     : Db 134 PSQGLLEDSKRTIEFEMPDANSTISRAAPAFASFPPLF--TGRMLKKEYSVLCKRCLAP 252
79
80     : Oy 182 RLHYIKKADENKGNFQKNTAFECSTHVPETAFIATVSYON-----HKTDLKTENMPFAK 234
81     : Db 253 RLKEMKQOE--SAAGSKVLARC--ETSSSESLRFKFWKGNELNRKNRKPQKILQKPF-- 307
82
83     : Db 235 GFRGSDIMELIHKSRMKSKEYEVPVRSYTRAKVNSHSPSSSESRALSTSLNLSQYQ-- 292
84     : Db 308 -----GSHLRIRKNSLADSDGSECMKCYIKSLDADSNANLTI--VESNATVSTIR--GTSLVLYK 362
85
86     : Oy 293 -----CEGNG-----VSPGDDLLPDPNXY 311
87     : Db 363 CAEKETFCVCNCEGCFYKWDLSNRYLCKCPNEF 397
88
89     : RESULT 8
90     : US-08-467-602-366
91     : Sequence 366, Application US/08467602C
92
93     : GENERAL INFORMATION:
94     : APPLICANT: Sklar, Robert
95     : APPLICANT: Marchionni, Mark
96     : APPLICANT: Gymer, David L FOR TREATING MUSCLE DISEASES AND
97     : TITLE OF INVENTION: DISORDERS
98     : FILE REFERENCE: 04585/028003
99     : CURRENT APPLICATION NUMBER: US/08/467, 602C
100    : CURRENT FILING DATE: 1995-06-06
101    : EARLIER FILING DATE: 1994-03-08
102    : EARLIER APPLICATION NUMBER: 08/059, 022
103    : EARLIER FILING DATE: 1993-05-06
104    : NUMBER OF SEQ ID NOS: 420
105    : SOFTWARE: FASTSEQ for Windows Version 4.0
106
107    : SEQ ID NO 366
108    : LENGTH: 456
109    : TYPE: PRT
110    : ORGANISM: Homo sapiens
111
112    : FEATURES:
113    : LOCATION: (34)
114    : OTHER INFORMATION: Xaa is any amino acid
115
116    : US-08-467-602-366
117
118    : Query Match
119    : Best Local Similarity 5.2%; Score 98.5; DB 4; Length 456;
120    : Best Local Similarity 23.0%; Pred. No. 0.045;
121    : Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18
122
123    : 25 DSKPRS-----ALGAPSKS-----PSSPOLAFTQOGMEGIVKFLIERELMLKTFHEVG 71
124    : Db DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
125
126    : 114 DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
127    : Db DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
128
129    : Oy 72 TELLITKAGRIKMPFSKYKVVYMGINPKRTKYLILMDIIVPADDHRYKFRKADNMKSVTQAKERP 131
130    : Db TELLITKAGRIKMPFSKYKVVYMGINPKRTKYLILMDIIVPADDHRYKFRKADNMKSVTQAKERP 131
131
132    : 168 -----NAVKAAGC-----KKDSILVLVL-----GTV-----CHPAF 193
133    : Db 132 P--GRL-----YVHPDPSATGAMKMRQVLVSQFKLKTNNHLDFFGHILNLMHKYOP 181
134    : Db 134 PSQGLLEDSKRTIEFEMPDANSTISRAAPAFASFPPLF--TGRMLKKEYSVLCKRCLAP 252
135
136    : Oy 182 RLHYIKKADENKGNFQKNTAFECSTHVPETAFIATVSYON-----HKTDLKTENMPFAK 234
137    : Db 253 RLKEMKQOE--SAAGSKVLARC--ETSSSESLRFKFWKGNELNRKNRKPQKILQKPF-- 307
138
139    : Db 235 GFRGSDIMELIHKSRMKSKEYEVPVRSYTRAKVNSHSPSSSESRALSTSLNLSQYQ-- 292
140    : Db 308 -----GSHLRIRKNSLADSDGSECMKCYIKSLDADSNANLTI--VESNATVSTIR--GTSLVLYK 362
141
142    : Oy 293 -----CEGNG-----VSPGDDLLPDPNXY 311
143    : Db 363 CAEKETFCVCNCEGCFYKWDLSNRYLCKCPNEF 397
144
145    : RESULT 8
146    : US-08-467-602-366
147    : Sequence 366, Application US/08467602C
148
149    : GENERAL INFORMATION:
150    : APPLICANT: Sklar, Robert
151    : APPLICANT: Marchionni, Mark
152    : APPLICANT: Gymer, David L FOR TREATING MUSCLE DISEASES AND
153    : TITLE OF INVENTION: DISORDERS
154    : FILE REFERENCE: 04585/028003
155    : CURRENT APPLICATION NUMBER: US/08/467, 602C
156    : CURRENT FILING DATE: 1995-06-06
157    : EARLIER FILING DATE: 1994-03-08
158    : EARLIER APPLICATION NUMBER: 08/059, 022
159    : EARLIER FILING DATE: 1993-05-06
160    : NUMBER OF SEQ ID NOS: 420
161    : SOFTWARE: FASTSEQ for Windows Version 4.0
162
163    : SEQ ID NO 366
164    : LENGTH: 456
165    : TYPE: PRT
166    : ORGANISM: Homo sapiens
167
168    : FEATURES:
169    : LOCATION: (34)
170    : OTHER INFORMATION: Xaa is any amino acid
171
172    : US-08-467-602-366
173
174    : Query Match
175    : Best Local Similarity 5.2%; Score 98.5; DB 4; Length 456;
176    : Best Local Similarity 23.0%; Pred. No. 0.045;
177    : Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18
178
179    : 25 DSKPRS-----ALGAPSKS-----PSSPOLAFTQOGMEGIVKFLIERELMLKTFHEVG 71
180    : Db DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
181
182    : 114 DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
183    : Db DREPPACGRRALGPRALGPEELLALANTVSPMLATVPSAGDEG-----EELRYLKVHGV 167
184
185    : Oy 72 TELLITKAGRIKMPFSKYKVVYMGINPKRTKYLILMDIIVPADDHRYKFRKADNMKSVTQAKERP 131
186    : Db TELLITKAGRIKMPFSKYKVVYMGINPKRTKYLILMDIIVPADDHRYKFRKADNMKSVTQAKERP 131
187
188    : 168 -----NAVKAAGC-----KKDSILVLVL-----GTV-----CHPAF 193
189    : Db 132 P--GRL-----YVHPDPSATGAMKMRQVLVSQFKLKT
```

```

Db      202 -----MAVKKAGT- KRDSLTIVRL-----GTW-----GHPAF 227
OY      133 P--GRL-----VYHDSPATGAMHMMROLDVSRQKLTJLNHHLDHPGHSILNMSMKYOP 181
      |||      |||      |||      |||      |||      |||      |||      |||
Db      228 PSCGRILKEDSNVITFMEEDNANSTSRAPAPRASPPLE-TGNIKLAKRSRYVLCRCALP 286
OY      182 RLHTIYAEDNNGGSKNATACTIVEPEPAIVATSYON-----IKITQLENNPFK 234
      |||      |||      |||      |||      |||      |||      |||      |||
Db      287 RLKEMKSDP-SAAGSKSLYLK-RTSSEISLRFKFKFNGNELNRKNKPNQITKDKP--- 341
OY      235 GFQSGDDELRLHRASRQSKETIYVPRSTVROKVASNHSPPSSERALSSTSSNLSQOTY-- 292
      |||      |||      |||      |||      |||      |||      |||      |||
Db      342 ---GKSELIRIKASLADSGEMCKVYSLGDNSASANVIT-VESNATSTST-CTSILVK 396
OY      293 -----CENG-----VSGPSODLLPPNPY 311
      |||      |||      |||      |||      |||      |||      |||      |||
Db      397 CAEKETPCVGVGEECPKWKDLSNPSRYLCAICGNNE 431
OY

```

```

1 RESULT 9
2 US-08-470-335-233
3 Sequence 233, Application US/08470335F
4 Patent No. 6147190
5 GENERAL INFORMATION:
6 APPLICANT: GOODEARL, ANDREW
7 APPLICANT: STROOBANT, PAUL
8 APPLICANT: MINCHETTI, LUISA
9 APPLICANT: WATERFIELD, MICHAEL
10 APPLICANT: MARCONIONI, MARK
11 APPLICANT: CHEN, MARIO S.
12 APPLICANT: HILES, IAN
13 TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
14 TITLE OF INVENTION: PREPARATION AND USE
15 FILE REFERENCE: 04585/00200B
16 CURRENT APPLICATION NUMBER: US/08/470,335F
17 CURRENT FILING DATE: 1995-06-06
18 EARLIER APPLICATION NUMBER: 08/036,555
19 EARLIER FILING DATE: 1993-03-24
20 NUMBER OF SEQ ID NOS: 252
21 SOFTWARE: FastSeq for Windows Version 4.0
22 SEQ ID NO 233
23 LENGTH: 601
24 TYPE: PRT
25 ORGANISM: Homo sapiens
26 FEATURE:
27 NAME/KEY: VARIANT
28 LOCATION: (601)...(601)
29 OTHER INFORMATION: Xaa is Arg.
30 US-08-470-335-233

```

Query Match 5.2%; Score 98.5; DB 4; Length 601;  
Best Local Similarity 23.0%; Pred. No. 0.07;  
Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18

Qy	25	DSKPEP-----ALGAPRS-----PSPQAPATQOQHEGIVLHEDELTMLFHEHG	71
Dd	114	DREPPAGRGALGPAAEPIILAAQGTVPAPVSGRG-----EAPVLYLVHQV---	167
Qy	72	TEMITACGRMRFPYVYKVTGTLGINKTKYITLMDIVPADHNRKFAQDKSVTGAARAP	131
Dd	168	-----MAVKAAGL---KQDSLVLVRL-----GTV-----GHPAF	193
Qy	132	P---GRL-----YVPSRPTAGAHMKRQVLYPOKAKILNHLDFGHTILNSHKKQRP	181
Dd	194	PCSGALHEDSRITFIFMERDANSTNAAPRAASFPLE---TGRNLKVEYSVLQKCALRP	252
Qy	182	RLHYKADENMGSGAKTATCTHYRPELAVIYTSYON-----HKTQLKLENNRPAPK	234
Dd	253	RLKEKESQF--SAAGSKVLVRC--ETISEYSLRFAKFKKNGELNRKKRQNKQILQKRP---	307
Qy	235	GFRSGDMELHRMSRMOSEKYEIVPVRSYRQKVASNIHSPSSERFALSTSSNLGSOY--	292
Dd	308	---GSGELRTINNAISLADSGYCKVYKISLQNDQASANITL---VESANSTSTLT---GTSHLYK	362

```

QY      293  -----CENG-----VSGCSQDLLPPPNPY 311
          | |
          : | | : | | :
Db      363  CAEKETFCVNGGECFMWKDLSNPSRYLCKCPNEF 397

```

RESULT 10  
US-08-467-602-323  
; Sequence 323, Application US/08467602C

```

1 TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND
2
3 FILE REFERENCE: 04585/028003
4
5 CURRENT APPLICATION NUMBER: US/08/467,602C
6
7 CURRENT FILING DATE: 1995-06-06
8
9 EARLIER APPLICATION NUMBER: 08/209,204
10
11 EARLIER FILING DATE: 1994-03-08
12
13 EARLIER APPLICATION NUMBER: 08/059,022
14
15 EARLIER FILING DATE: 1993-05-06
16
17 NUMBER OF SEQ ID NOS: 420
18
19 SOFTWARE: FASTSEQ for Windows Version 4.0
20
21 SEQ ID NO 323
22
23 LENGTH: 601
24
25 TYPE: prt
26
27 ORGANISM: Homo sapiens
28
29 FEATURE:
30
31 NAME/KEY: VARIANT
32
33 LOCATION: (601)...(601)
34
35 OTHER INFORMATION: Xaa is Arg or absent.
36
37 US-08-467,602-323

```

Query Match	5.28; Score 98.5; DB 4; Length 601;
Query Match	5.28; Score 98.5; DB 4; Length 601;

Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18;

[illegible]

```
QY      18Z RLHIKADENNGFGSKNTAFCTHVEFPETARIIVTSYON-----HKTTOLKIENNPFAPK 234
```

QY 235 GFPSODMELHBSR0KXEYPVPRSTVOXKVASNNSPSSSALSTSSNGSOYO-- 292  
 : : : : :  
 Db 308 ---GKEBLRIKMSLAOSGEYMKCVISLONDSANNTI-VESNATISTTT-OTSHLYK 362  
 QY 293 -----CENG-----VSOPSDLLPPPV 311  
 : : : : :  
 Db 363 CAEKETFCVNGECFTMYKDLNSPRTLCGRNEF 397

RESULT 11  
 US-08-470-335-236  
 Sequence 236, Application US/08470335F  
 Patent No. 6147190  
 GENERAL INFORMATION:  
 APPLICANT: GOODDEARL, ANDREW  
 APPLICANT: STROOBANT, PAUL  
 APPLICANT: MINGCHETTI, LUISA  
 APPLICANT: WATERFIELD, MICHAEL

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.ra1

Page 6

```

APPLICANT: MARCHIONNI, MARK
APPLICANT: CHEN, MARIO S.
TITLE OF INVENTION: CLINAL MITOGENIC FACTORS, THEIR
FILE REFERENCE: 04585/0200B
CURRENT FILING DATE: 1995-06-06
EARLIER FILING DATE: 1993-03-24
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 236
LENGTH: 610
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (610)...(610)
OTHER INFORMATION: Xaa is Arg.
US-08-467-602-332

Query Match      5.2%: Score 98.5; DB 4; Length 610;
Best Local Similarity 23.0%: Pred. No. 0.072;
Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18;

OY 25 DSKPES---ALGAPSKS-----PSSPOAFTQOQEGIKYFLHRELMKHEHG 71
DB 114 DREPPAGPRLGPPAEPELLAANGTVSNPFAVPSAGEFG---EAPFLYKVIQV- 167
OY 72 TEMITKAGRRMPSKYKRVYGLNFKTKYLLMDIYVADDDHRYKFAADKKWSTGAKAPM 131
DB 168 -----NAVAKAGL--KQDSLTVRL-----GTW-----GHPAF 193
OY 132 P--GRL-----YHDPSPATGAMHMQOLVSOKLKTNNHLPFGHILINSMHKYO 181
DB 194 PSCGRLEKEDSRITFMEPDANSTRAPAFRAFPPELE--TGRNLKKEVSVLCRCKALP 252
OY 182 RLHIYKADENNGFSKNTAFCTHVPPEFAETAVTSYON-----HKITOLKIENNPF 234
DB 235 GFPGSDMEJLHMSQSKETPVVPRSTVROKVASNHSPFSSSRALSTSSNLGSOYO-- 292
OY 308 ---GSELRINKASLADSGEYKCKYKSLANDNSANITV--VESNATYSTT--GTSHLYK 362
DB 293 -----CENG-----VSGPSQDLPPNNY 311
OY 363 CAEKERTFCVNGCECFWADLSPSRYLCKCNEF 397
DB

RESULT 12
US-08-467-602-332
Sequence 332, Application US/08467602C
Patent No. 644642
GENERAL INFORMATION:
APPLICANT: Sklar, Robert
APPLICANT: Marchionni, Mark
APPLICANT: Gayme, David I.
TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND
FILE REFERENCE: 04585/028003
CURRENT APPLICATION NUMBER: US/08/467,602C
EARLIER FILING DATE: 1994-03-08
EARLIER APPLICATION NUMBER: 08/209,204
EARLIER FILING DATE: 1993-05-06
NUMBER OF SEQ ID NOS: 420
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 236
LENGTH: 610
TYPE: PRT
US-08-467-602-332
```

```

ORGANISM: Homo sapiens
NAME/KEY: VARIANT
LOCATION: (610)...(610)
OTHER INFORMATION: Xaa is Arg or absent.
US-08-467-602-332

Query Match      5.2%: Score 98.5; DB 4; Length 610;
Best Local Similarity 23.0%: Pred. No. 0.072;
Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18;

OY 25 DSKPES---ALGAPSKS-----PSSPOAFTQOQEGIKYFLHRELMKHEHG 71
DB 114 DREPPAGPRLGPPAEPELLAANGTVSNPFAVPSAGEFG---EAPFLYKVIQV- 167
OY 72 TEMITKAGRRMPSKYKRVYGLNFKTKYLLMDIYVADDDHRYKFAADKKWSTGAKAPM 131
DB 168 -----NAVAKAGL--KQDSLTVRL-----GTW-----GHPAF 193
OY 132 P--GRL-----YHDPSPATGAMHMQOLVSOKLKTNNHLPFGHILINSMHKYO 181
DB 194 PSCGRLEKEDSRITFMEPDANSTRAPAFRAFPPELE--TGRNLKKEVSVLCRCKALP 252
OY 182 RLHIYKADENNGFSKNTAFCTHVPPEFAETAVTSYON-----HKITOLKIENNPF 234
DB 235 GFPGSDMEJLHMSQSKETPVVPRSTVROKVASNHSPFSSSRALSTSSNLGSOYO-- 292
OY 308 ---GSELRINKASLADSGEYKCKYKSLANDNSANITV--VESNATYSTT--GTSHLYK 362
DB 293 -----CENG-----VSGPSQDLPPNNY 311
OY 363 CAEKERTFCVNGCECFWADLSPSRYLCKCNEF 397
DB

RESULT 13
US-08-467-602-365
Sequence 365, Application US/08467602C
Patent No. 644642
GENERAL INFORMATION:
APPLICANT: Sklar, Robert
APPLICANT: Marchionni, Mark
APPLICANT: Gayme, David I.
TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND
FILE REFERENCE: 04585/028003
CURRENT APPLICATION NUMBER: US/08/467,602C
EARLIER FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: 08/209,204
EARLIER FILING DATE: 1994-03-08
EARLIER APPLICATION NUMBER: 08/059,022
NUMBER OF SEQ ID NOS: 420
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 235
LENGTH: 615
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (34)...(34)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: VARIANT
LOCATION: (635)...(635)
OTHER INFORMATION: Xaa is Arg or absent.
US-08-467-602-365

Query Match      5.2%: Score 98.5; DB 4; Length 635;
Best Local Similarity 23.0%: Pred. No. 0.076;
Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18;
```

```

OY 25 DSKPRS-----ALGAPSKS-----PSPQAAFTQOGMEGICYFLHERELMLKFEHVG 71
Db 148 DREPPAAGPRALGPPAEPELLAANGVPSMPTAPVPSAGEG-----EAPRYIVYHQV- 201
OY 72 TSMITTAGRRMFPSYKVKVYGLNPRTKTYILLMDIYPADDRKRFKADKKSVYGAEPAA 131
Db 202 -----NAVYAGGL--KKDSLLTYRL-----GTYW-----GHPAF 227
OY 132 P--GRL-----YVHPDSPATGAMHMQOLYSFOKLITJNNHLDPFCHITILNSMKNQOP 181
Db 228 PSCGKLKEOSRYIFEMEPDANSTSRAPAFRASPPLLE-TGRNLKKEVSRYLCKRCALPP 286
OY 182 RLHYKADENNFGSKNTAFCTHVPEPTAFIAYTSYON-----HKITOLKIENPFPAK 234
Db 287 RLKEKKSQE-SAAGSKIVLYRC-ETISEVSYSLRFKFWKNGNELMRKKNPOINKLOKPP--- 341
OY 235 GFRGSDMELHMRSMQSKKEYPVVPSTYRQKVASNHSPFSESRALSTSSNLGSOYO-- 292
Db 342 ---GKSELIRINKASLADSGEYCKVSKLGDNSANITTI-VESNATSTSTT-GTSHLYK 396
OY 293 -----CENG-----VSGSODLLPPPNPY 311
Db 397 CAEKREKTFVNGGCEFWKDLNSBRYLCKCPNEF 431

RESULT 14
US-08-467-602-374
Sequence 374, Application US/08467602C
Patent No. 6444642
GENERAL INFORMATION:
APPLICANT: SKIAT, Robert
APPLICANT: MARCHIONNI, MARK
APPLICANT: GARYME, DAVID I.
TITLE OF INVENTION: METHODS FOR TREATING MUSCLE DISEASES AND
FILE REFERENCE: 04585/002008
CURRENT APPLICATION NUMBER: US/08/467,602C
EARLIER FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: 08/209,204
EARLIER FILING DATE: 1994-03-08
EARLIER APPLICATION NUMBER: 08/059,022
NUMBER OF SEQ ID NOS: 420
SOFTWARE: FASTSQ for Windows Version 4.0
SEQ ID NO 374
LENGTH: 644
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (34)...(34)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: VARIANT
LOCATION: (644)...(644)
OTHER INFORMATION: Xaa is Arg or absent.
US-08-467-602-374

Query Match 5.2%: Score 98.5; DB 4; Length 644;
Best Local Similarity 23.0%: Pred. No. 0.078;
Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18;

OY 25 DSKPRS-----ALGAPSKS-----PSPQAAFTQOGMEGICYFLHERELMLKFEHVG 71
Db 148 DREPPAAGPRALGPPAEPELLAANGVPSMPTAPVPSAGEG-----EAPRYIVYHQV- 201
OY 72 TSMITTAGRRMFPSYKVKVYGLNPRTKTYILLMDIYPADDRKRFKADKKSVYGAEPAA 131
Db 202 -----NAVYAGGL--KKDSLLTYRL-----GTYW-----GHPAF 227
OY 132 P--GRL-----YVHPDSPATGAMHMQOLYSFOKLITJNNHLDPFCHITILNSMKNQOP 181
Db 228 PSCGKLKEOSRYIFEMEPDANSTSRAPAFRASPPLLE-TGRNLKKEVSRYLCKRCALPP 286
OY 293 -----CENG-----VSGSODLLPPPNPY 311
Db 397 CAEKREKTFVNGGCEFWKDLNSBRYLCKCPNEF 431

```

```

OY 182 RLHYKADENNFGSKNTAFCTHVPEPTAFIAYTSYON-----HKITOLKIENPFPAK 234
Db 287 RLKEKKSQE-SAAGSKIVLYRC-ETISEVSYSLRFKFWKNGNELMRKKNPOINKLOKPP--- 341
OY 235 GFRGSDMELHMRSMQSKKEYPVVPSTYRQKVASNHSPFSESRALSTSSNLGSOYO-- 292
Db 342 ---GKSELIRINKASLADSGEYCKVSKLGDNSANITTI-VESNATSTSTT-GTSHLYK 396
OY 293 -----CENG-----VSGSODLLPPPNPY 311
Db 397 CAEKREKTFVNGGCEFWKDLNSBRYLCKCPNEF 431

RESULT 15
US-08-470-335-234
Sequence 234, Application US/08470335F
Patent No. 6147190
GENERAL INFORMATION:
APPLICANT: GOODEARL, ANDREW
APPLICANT: STROOBANT, PAUL
APPLICANT: MINGHETTI, LUISA
APPLICANT: WATERFIELD, MICHAEL
APPLICANT: MARCHIONNI, MARK
APPLICANT: CHEN, MARIO S.
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
TITLE OF INVENTION: PREPARATION AND USE
FILE REFERENCE: 04585/002008
CURRENT APPLICATION NUMBER: US/08/470,335F
EARLIER FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: 08/036,555
EARLIER FILING DATE: 1993-03-24
NUMBER OF SEQ ID NOS: 252
SOFTWARE: FASTSQ for Windows Version 4.0
SEQ ID NO 234
LENGTH: 818
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (34)...(34)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: VARIANT
LOCATION: (644)...(644)
OTHER INFORMATION: Xaa is Arg or absent.
US-08-470-335-234

Query Match 5.2%: Score 98.5; DB 4; Length 818;
Best Local Similarity 23.0%: Pred. No. 0.12;
Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18;

OY 25 DSKPRS-----ALGAPSKS-----PSPQAAFTQOGMEGICYFLHERELMLKFEHVG 71
Db 114 DREPPAAGPRALGPPAEPELLAANGVPSMPTAPVPSAGEG-----EAPRYIVYHQV- 167
OY 72 TSMITTAGRRMFPSYKVKVYGLNPRTKTYILLMDIYPADDRKRFKADKKSVYGAEPAA 131
Db 168 -----NAVYAGGL--KKDSLLTYRL-----GTYW-----GHPAF 193
OY 132 P--GRL-----YVHPDSPATGAMHMQOLYSFOKLITJNNHLDPFCHITILNSMKNQOP 181
Db 194 PSCGKLKEOSRYIFEMEPDANSTSRAPAFRASPPLLE-TGRNLKKEVSRYLCKRCALPP 252
OY 182 RLHYKADENNFGSKNTAFCTHVPEPTAFIAYTSYON-----HKITOLKIENPFPAK 234
Db 253 RLKEKKSQE-SAAGSKIVLYRC-ETISEVSYSLRFKFWKNGNELMRKKNPOINKLOKPP--- 307
OY 235 GFRGSDMELHMRSMQSKKEYPVVPSTYRQKVASNHSPFSESRALSTSSNLGSOYO-- 292
Db 308 ---GKSELIRINKASLADSGEYCKVSKLGDNSANITTI-VESNATSTSTT-GTSHLYK 362
OY 293 -----CENG-----VSGSODLLPPPNPY 311
Db 363 CAEKREKTFVNGGCEFWKDLNSBRYLCKCPNEF 397

```

Search completed: June 17, 2003, 09:49:47  
Job time : 16 secs

**THIS PAGE BLANK (uspto)**





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:49:08 ; Search time 21 Seconds  
(without alignments) 1776.837 Million cell updates/sec

Title: US-10-027-059a-1  
Perfect score: 1889  
Sequence: 1 MADADGFGIAHPLPDKA.....CDHPMSICFSLFSLGNG 349

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:\*  
1: /cgn2\_6/prodata/1/pubpa/US06\_NEW\_PUB.pep.\*  
2: /cgn2\_6/prodata/1/pubpa/PCIT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/1/pubpa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/1/pubpa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubpa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/1/pubpa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/1/pubpa/US07\_PUBCOMB.pep.\*  
8: /cgn2\_6/prodata/1/pubpa/US07\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/1/pubpa/US07\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/1/pubpa/US07\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/1/pubpa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/prodata/1/pubpa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/prodata/1/pubpa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/prodata/1/pubpa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1889	100.0	349	US-10-027-059a-1	Sequence 1, App1
2	1757	93.5	518	US-10-027-059a-3	Sequence 3, App1
3	535	28.3	530	US-10-008-264a-4	Sequence 4, App1
4	528.5	28.0	535	US-10-008-264a-2	Sequence 2, App1
5	517.5	27.4	365	US-09-925-300-1676	Sequence 1676, App
6	158	8.4	57	US-09-864-761-33963	Sequence 33963, A
7	99	5.2	24	US-09-983-802-522	Sequence 522, App
8	98.5	5.2	418	US-09-946-807-3	Sequence 3, App1
9	98.5	5.2	418	US-09-795-668-3	Sequence 3, App1
10	98.5	5.2	418	US-09-795-668-3	Sequence 3, App1
11	94	5.0	30	US-09-864-122-20	Sequence 39230, A
12	93	4.9	320	US-09-854-171-20	Sequence 20, App1
13	92.5	4.9	1252	US-09-908-193-19	Sequence 19, App1
14	92.5	4.9	1252	US-09-908-193-20	Sequence 20, App1
15	91.5	4.8	588	US-09-801-368-46	Sequence 46, App1
16	91.5	4.8	1252	US-09-908-193-21	Sequence 21, App1
17	89	4.7	1115	US-09-794-997a-2	Sequence 2, App1
18	89	4.7	1115	US-10-174-590-440	Sequence 440, App
19	89	4.7	1115	US-10-176-758-440	Sequence 440, App

20	89	4.7	1115	US-10-175-737-440	Sequence 440, App
21	89	4.7	1115	US-10-173-706-440	Sequence 440, App
22	89	4.7	1115	US-10-175-738-440	Sequence 440, App
23	89	4.7	1115	US-10-175-752-440	Sequence 440, App
24	89	4.7	1115	US-10-176-482-440	Sequence 440, App
25	89	4.7	1115	US-10-176-757-440	Sequence 440, App
26	89	4.7	1115	US-10-176-913-440	Sequence 440, App
27	89	4.7	1115	US-10-180-552-440	Sequence 440, App
28	89	4.7	1115	US-10-180-557-440	Sequence 440, App
29	89	4.7	1115	US-10-173-700-440	Sequence 440, App
30	89	4.7	1115	US-10-174-572-440	Sequence 440, App
31	89	4.7	1115	US-10-174-579-440	Sequence 440, App
32	89	4.7	1115	US-10-174-582-440	Sequence 440, App
33	89	4.7	1115	US-10-175-739-440	Sequence 440, App
34	89	4.7	1115	US-10-175-740-440	Sequence 440, App
35	89	4.7	1115	US-10-175-743-440	Sequence 440, App
36	89	4.7	1115	US-10-176-488-440	Sequence 440, App
37	89	4.7	1115	US-10-176-488-440	Sequence 440, App
38	89	4.7	1115	US-10-176-492-440	Sequence 440, App
39	89	4.7	1115	US-10-176-747-440	Sequence 440, App
40	89	4.7	1115	US-10-176-750-440	Sequence 440, App
41	89	4.7	1115	US-10-176-985-440	Sequence 440, App
42	89	4.7	1115	US-10-176-987-440	Sequence 440, App
43	89	4.7	1115	US-10-176-991-440	Sequence 440, App
44	89	4.7	1115	US-10-176-992-440	Sequence 440, App
45	89	4.7	1115	US-10-176-993-440	Sequence 440, App

#### ALIGNMENTS

RESULT 1  
US-10-027-059a-1  
Sequence 1, Application US/1007059A  
Patent No. US20020165382A1  
APPLICANT: BASCON, Craig  
TITLE OF INVENTION: Transcription Factors that Regulate No. US20020165382A1 and  
FILE REFERENCE: 955-1AP  
CURRENT FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1  
LENGTH: 349  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-027-059a-1

Query Match: 100.0%; Score 1889; DB 9; Length 349;  
Best Local Similarity 100.0%; Pred. No. 3.8e-168;  
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MADADGFGIAHPLPDKA.....CDHPMSICFSLFSLGNG 60  
1 MADADGFGIAHPLPDKA.....CDHPMSICFSLFSLGNG 60  
61 RELMLAFHEVGTETMTTAKGRARFPEYKVVYVGTLPKTKYLLMDIYVPAADHRRKFAADK 120  
61 RELMLAFHEVGTETMTTAKGRARFPEYKVVYVGTLPKTKYLLMDIYVPAADHRRKFAADK 120  
61 RELMLAFHEVGTETMTTAKGRARFPEYKVVYVGTLPKTKYLLMDIYVPAADHRRKFAADK 120  
121 WSVTGAEPAPMPGRLVHPDPSPTAGAHMQLVSPQKLTJNNHLDPFCHITLNSHMKYQ 180  
121 WSVTGAEPAPMPGRLVHPDPSPTAGAHMQLVSPQKLTJNNHLDPFCHITLNSHMKYQ 180  
121 WSVTGAEPAPMPGRLVHPDPSPTAGAHMQLVSPQKLTJNNHLDPFCHITLNSHMKYQ 180  
181 PELHAIKADENNGSGSKTACFCHVPEPTAFVLTYSYONKIKITOLKLENNPAGFGNSD 240  
181 PELHAIKADENNGSGSKTACFCHVPEPTAFVLTYSYONKIKITOLKLENNPAGFGNSD 240  
241 DMELHAKMSRMSKEYEYVPRSTYRQKVASNHSPFSSESRALSTSSNGSYOCENYSGSP 300  
241 DMELHAKMSRMSKEYEYVPRSTYRQKVASNHSPFSSESRALSTSSNGSYOCENYSGSP 300

```
OY 301 SODLLPPNPVPLPOHSHQVICTKRRKGGCDHPWSCICFLSTLSTLNGC 349
DB 301 SODLLPPNPVPLPOHSHQVICTKRRKGGCDHPWSCICFLSTLSTLNGC 349

RESULT 2
US-10-027-059a-3
Sequence 3, Application US/10027059a
Patent No. US20020165382A1
GENERAL INFORMATION: Ctrial
TITLE OF INVENTION: Transcription Factors that Regulate No. US20020165382A1mal and Mz
FILE REFERENCE: 955-122
CURRENT FILING DATE: US/10/027.059a
NUMBER OF SBO ID NOS: 15
SOFTWARE: PatentIn version 3.0
SBO ID NO 518
TYPE: PRT
ORGANISM: Homo sapiens
US-10-027-059a-3

Query Match 93.5%; Score 1767; DB 9; Length 518;
Best Local Similarity 97.9%; Pred. No. 1.7e-1156;
Matches 331; Conservative 1; Mismatches 2; Indels 4; Gaps 1:

OY 1 MDADDEGFLAHPLEPDANKDPCSPSPSALGAPSKSPSPSPQAFTQOCMEGIXYFLHE 60
DB 1 MDADDEGFLAHPLEPDANKDPCSPSPSALGAPSKSPSPSPQAFTQOCMEGIXYFLHE 60

OY 61 RELMLAFHEVGTETITKAGRMFSPYKVKVYGLNFKTKYLLMDIYPADDRHFKFADNK 120
DB 61 RELMLAFHEVGTETITKAGRMFSPYKVKVYGLNFKTKYLLMDIYPADDRHFKFADNK 120

OY 121 WSYTGKAPFAMPGALYVHPSDPATGAMRNOLVSFOKLTILNNHDPFGHITLMSMKYQ 180
DB 121 WSYTGKAPFAMPGALYVHPSDPATGAMRNOLVSFOKLTILNNHDPFGHITLMSMKYQ 180

OY 181 PRHLYVADENNGFSKNTAFCHVPEPAFAVATSYONKHTIOLKIENNPFAGFNGSD 240
DB 181 PRHLYVADENNGFSKNTAFCHVPEPAFAVATSYONKHTIOLKIENNPFAGFNGSD 240

OY 241 DMLHHRMSKQKEYVVPYPSYRQYASMSPPSESRALSTSENCGOYPCANGYSGP 300
DB 241 DMLHHRMSKQKEYVVPYPSYRQYASMSPPSESRALSTSENCGOYPCANGYSGP 300

OY 301 SODLLPPNPVPLPOHSHQVICTKRRKGGCDHPWSCICFLSTLSTLNGC 349
DB 301 SODLLPPNPVPLPOHSHQVICTKRRKGGCDHPWSCICFLSTLSTLNGC 349

RESULT 3
US-10-008-264a-4
Sequence 4, Application US/10008264a
Publication No. US20030104528A1
GENERAL INFORMATION:
APPLICANT: Glanther, Laurie H.
APPLICANT: Szabo, Susanne, J.
FILE REFERENCE: H01-040CP
CURRENT FILING DATE: US/10/008.264a
PRIOR APPLICATION NUMBER: PCT/US00/15345
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/137085
PRIOR FILING DATE: 1999-06-02
SOFTWARE: PatentIn Ver. 2.0
SBO ID NO 4
TYPE: PRT
ORGANISM: Mus musculus
```

```
US-10-008-264a-4
Query Match 28.3%; Score 535; DB 9; Length 530;
Best Local Similarity 37.7%; Pred. No. 1.3e-43;
Matches 147; Conservative 40; Mismatches 97; Indels 134; Gaps 15:

OY 16 EPKAKDLPDCKSPESALGAP-----SKSPSPQAFTQOCMEGIXYFLHE 49
DB 36 EPKAKDLPDCKSPESALGAP-----SKSPSPQAFTQOCMEGIXYFLHE 49

OY 50 GMEG-----IKVFLHRELMLAFHEVGTETIT 74
DB 97 GADGYPPIODGYAPDPAGALYVHPSDPATGAMRNOLVSFOKLTILNNHDPFGHITLMSMKYQ 156

OY 75 IITKAGRMFSPYKVKVYGLNFKTKYLLMDIYPADDRHFKFADNKFAMFNGKAPFAG- 133
DB 157 IITKAGRMFSPYKVKVYGLNFKTKYLLMDIYPADDRHFKFADNKFAMFNGKAPFAG- 216

OY 134 RLIVHPSDPATGAMRNOLVSFOKLTILNNHDPFGHITLMSMKYQPRHLIVYADE 190
DB 217 RLIVHPSDPATGAMRNOLVSFOKLTILNNHDPFGHITLMSMKYQPRHLIVYADE 276

OY 191 NNGFSKNTAFCHVPEPAFAVATSYONKHTIOLKIENNPFAGFNGSD----- 237
DB 277 -----GEPEAKCSASNTHTFPOETOFLAVATSYONKHTIOLKIENNPFAGFNGSD 331

OY 238 -----GSDMLHHRMSKQKEYV-----VPSTVNRK--- 266
DB 332 ASVDTYSVSPGNCOLLGDP-----FSPULSNDGYEVPSNFTDLPQADMIISQPYWL 386

OY 267 -VASNHSPPSESRALSTSENCGOYPCANGYSGP-----SODLLPPNPVPLPOH 317
DB 387 GTPREHS-YEAFRAVSMKPTL-----LPSAPGPTVYKQDVLAVGACGAPVAPQY 437

RESULT 4
US-10-008-264a-2
Sequence 2, Application US/10008264a
Publication No. US20030104528A1
GENERAL INFORMATION:
APPLICANT: Glanther, Laurie H.
APPLICANT: Szabo, Susanne, J.
FILE REFERENCE: H01-040CP
CURRENT FILING DATE: US/10/008.264a
PRIOR APPLICATION NUMBER: PCT/US00/15345
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/137085
PRIOR FILING DATE: 1999-06-02
NUMBER OF SBO ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SBO ID NO 2
TYPE: PRT
ORGANISM: Homo sapiens
US-10-008-264a-2

Query Match 28.0%; Score 528.5; DB 9; Length 535;
Best Local Similarity 37.7%; Pred. No. 7.8e-41;
Matches 140; Conservative 41; Mismatches 115; Indels 75; Gaps 12:

OY 3 DAFD-----GGLAHPLEPDANKDPCSPSPSALGAPSKSPSPQAFTQOC-----NE 52
DB 44 DAFDNRGSGSLSPY-RGDAVLP--APPSNFLGAVAPPROAMGFPACGESEFPAPDAE 100

OY 53 G-----IKVFLHRELMLAFHEVGTETIT 77
DB 101 GYDPEGYAAPDPAGALYVHPSDPATGAMRNOLVSFOKLTILNNHDPFGHITLMSMKYQ 160

OY 78 KAGRMFSPYKVKVYGLNFKTKYLLMDIYPADDRHFKFADNKFAMFNGKAPFAG-RLX 136
DB 161 KGRMFPELSTYVAGLEPSTYHRYMFDVVLVDHHRMSKQKEYVVPYPSYRQYASMSPPSESRAL 220
```

```
QY 137 VHPDSPATGAMHNRQVLSFQKLTNTN--HILDPGHIILNSMKRYOPRLHYKADENNG 193
DB 221 VHPDSPNGAMHNRQVLSFQKLTNTNKKASNNVTQIVLQSLKROPRLHYKADENNG 280
QY 194 FGSKATATACATY--PPEATATVTSYONKRTKQKLTENNPFKAGSGSDDELEHRSNMQ 251
DB 281 EACN-ASNNHIFTEOFTQVAVYONAEITQKLTENNPFKAGSGSDDELEHRSNMQ 335
QY 252 SKREYVYVRESTRYOKVANSNPSFSSSRALSTSSNLSQYOCEN---GVSGSDQLLP 307
DB 336 DTISPSPGNCOFGSDHYSPL-----LPNQYVPSRFYPLDQAKDVYPO 383
QY 308 PNFYPLPOENS 318
DB 384 AYLWLAPEHDS 394

RESULT 5
US-09-925-300-1676
: Sequence 1676, Application US/09925300
: Patent No. US20020151681A1
: GENERAL INFORMATION:
: APPLICANT: Steve Ruben,
: APPLICANT: Craig Rosen,
: TITLE OF INVENTION: Nucleic Acids, proteins and Antibodies
: FILE REFERENCE: PA101
: CURRENT APPLICATION NUMBER: US/09/925, 300
: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05988
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124, 270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1690
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO: 1676
: LENGTH: 365
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (47)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (220)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1676

Query Match 27.4%, Score 517.5, DB 10, Length 365;
Best Similarity 45.9%, Pred. No. 4,7e-40;
Matches 112, Conservative 33, Mismatches 74, Indels 25, Gaps 4;

QY 83 MFSKYVYVNGVGLPKTYITLMDYPRDHRKTKADKKSYTGAGAPRGLYVDPSP 142
DB 4 MFPFVYVNGVGLPKTYITLMDYPRDHRKTKADKKSYTGAGAPRGLYVDPSP 63
QY 143 ATGAMHMOAYSPQKLTNTNHLDPG-----HTTNSMKRYOPR 182
DB 64 ATGEQMSKVVYFHLKLTNTNSDKHFTLARPDDAHYMQGVSGFTQILNSMKRYOPR 123
QY 183 LHVADNNGSGSNATPCTVPPERPAFLVTSYONKRTKQKLTENNPFKAGSGSDM 242
DB 124 FHVIRA--NDILKLYSTFRYTLPEREFIAVTAQNDKLTQKLTENNPFKAGSDTNG 181
QY 243 ELHRNSRQKQKEYPVPRSTVYOKVANSNPSFSSSRALSTSSNLSQYOCENGVSGSO 302
DB 182 RRRKRRQTLQSMRYFDERHKKENGTSDS--SSQAAENFKQAQASPAATVQTSN-LK 238
QY 303 DLCP 306
DB 239 DLCP 242
```

```
RESULT 6
US-09-864-761-33963
: Sequence 33963, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aecomica-x-1
: CURRENT APPLICATION NUMBER: US/09/864, 761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180, 312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207, 456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632, 366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263, 6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236, 359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234, 687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608, 408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774, 203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 48117
: SOFTWARE: Jannus Sequence Listing Engine vers. 1.1
: SEQ ID NO: 33963
: LENGTH: 57
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: MAP TO AC007076.2
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
: OTHER INFORMATION: EXPRESSED IN HEPLIO, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN HEPT, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN FETAL SIGNAL = 2.4
: OTHER INFORMATION: EST HUMAN HIT: BE68028.1, EVALUE 5.00e-17
: OTHER INFORMATION: SWISSPROT HIT: Q94890, EVALUE 2.00e-19
US-09-864-761-33963

Query Match 8.4%, Score 158, DB 10, Length 57;
```

Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rapb

Page 4

RESULT: 7  
 Sequence 963-802-522  
 Publication No. US20030022185A1  
 GENERAL INFORMATION:  
 APPLICANT: Fischer et al.  
 TITLE OF INVENTION: 1.33 Human Secreted Proteins  
 CURRENT APPLICATION NUMBER: US/09/963\_802  
 PRIOR FILING DATE: 2001-10-25  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/327,357  
 PRIOR FILING DATE: 1999-01-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/12684  
 PRIOR FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,926  
 PRIOR FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,793  
 PRIOR FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,925  
 PRIOR FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,929  
 PRIOR FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,823  
 PRIOR FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,732  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,931  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,733  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,795  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,919  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,722  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,733  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,948  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,919  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,953  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,950  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,964  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,360  
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18  
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,684

```

OY      214 VTSYONKRTITOLKTIENPFANR 237
Db      1 VTSYONKRTITOLKTIENPFANR 24

RESULT 8
US-09-946-807-3
Sequence 3, Application US/09946807
GENETIC INFORMATION
APPLICANT: Stefansson, Helmi
APPLICANT: Steinhorstodottir, Valgeirur
APPLICANT: Gulcher, Jeffrey R.
TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
PRIORITY REFERENCE NUMBER: US/09/946,807
CURRENT FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: US/09/795,668
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US/09/515,716
PRIOR FILING DATE: 2000-02-28
PRIORITY REFERENCE NUMBER: US/09/515,716
SOFTWARE: FASTSD for Windows Version 4.0
SEQ ID NO 3
LENGTH: 418
TYPE: PART
ORGANISM: Homo sapiens
US-09-946-807-3
Query Match
Best Local Similarity 5.2%; Score 98.5; DB 9; Length 418;
Matches 77; Conservative 40; Mismatches 119; Indels 99; Gaps 18
OY      25 DSKPSS---ALCAPSKS-----PSSDQAFPOOQMCITKFLAEELHILKTEY 71
Db      11 DREPPACPALGPAPAEELLIANGVTSPMFAVPSAAGERG---EAPLVKAYOV- 163
OY      72 TEMITKAGRRFSSKRYKVGANGKTKYLLIMQIVADDRKFKADMKKSYSTCKEAPM 169
Db      11 YAVYAGAGI---KNDQLTVRL-----GTV-----GHPAF 199
OY      132 P---GGL-----NAPPSPACAKMMKQVLSQTKLKTNNHLDPRGHTILNSMKYOP 181
Db      190 PSCGLTEKEDSRITVFPEMDPANSISAPAPAEPSPLE---TORLAKKEVSIVCKKCALPP 248
OY      182 RLIVKAKDENKQFSSKNAFCFTVIEPFAIVSYQN-----HKITOLKTIENPFAR 234
Db      249 RLKKEKQSE--SNASQKVLKRC--ETSSSTSLKFKKFNQONELKRNKKKQVNIQAKAF--- 303

```



```

        prior APPLICATION NUMBER: PCT/US01/00668
        prior FILING DATE: 2001-01-30
        prior APPLICATION NUMBER: PCT/US01/00663
        prior FILING DATE: 2001-01-30
        prior APPLICATION NUMBER: PCT/US01/00662
        prior FILING DATE: 2001-01-30
        prior APPLICATION NUMBER: PCT/US01/00661
        prior FILING DATE: 2001-01-30
        prior APPLICATION NUMBER: PCT/US01/00670
        prior FILING DATE: 2001-01-30
        prior APPLICATION NUMBER: US 60/234,687
        prior FILING DATE: 2000-06-30
        prior APPLICATION NUMBER: US 09/608,408
        prior FILING DATE: 2000-06-30
        prior APPLICATION NUMBER: US 09/774,203
        prior FILING DATE: 2001-01-29
        NUMBER OF SEQ ID NOS: 49117
        SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
        SEQ ID NO: 39230
        LENGTH: 30
        TYPE: prt
        ORGANISM: Homo sapiens
FEATURES:
    OTHER INFORMATION: MAP TO AC005826.1
    OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
    OTHER INFORMATION: EXPRESSED IN HEPA-MARGON, SIGNAL = 2.0
    OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
    OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
    OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
    OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4
    OTHER INFORMATION: EST=UDM HRT: AM66610.1, EVALDE 7.00e-05
    OTHER INFORMATION: SWISSPROT HIT: Q90MK3, EVALDE 1.00e-11
US-09-854-122-20
Query Match          5.0%: Score 94; DB 10: Length 30;
Best Local Similarity 63.3%: Pred. No. 0.044;
Matches 19: Consecutive 3; Mismatches 8; Indels 0; Gaps 0.
OY      51 MCGIVTLERGGTKRFGNGTMMITKRG 80
Db      1 MAKIACSTELKEMLDKPHECZTEMLIKRSQ 30
RESULT 12
US-09-854-122-20
Sequence 40: N6010015988A1
Reference 40: US09854122
GENERAL INFORMATION:
APPLICANT: SMITH, ROBERT
APPLICANT: ALBERTE, RANDALL S.
TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
FILE REFERENCE: PHA-007.01
CURRENT APPLICATION NUMBER: US/09/854,122
PRIORITY CLAIMING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 60/202,529
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 20
LENGTH: 320
TYPE: part
OTHER INFORMATION: Flaveria bidentatis
US-09-854-122-20
Query Match          4.9%: Score 93; DB 10: Length 320;
Best Local Similarity 23.2%: Pred. No. 1.7;
Matches 60: Consecutive 28; Mismatches 97; Indels 74; Gaps 13.
OY      114 YRFADNKSVTGAKEMFGR--LIVVPS-----PATGHHNHLVSFOKLKLNHNL 165
Db      33 YTYDD-PMDGNMGICALIAGSQFARPDVYLSPFSQPTMTALA---VAIVTRKPF 88

```

QY 166 DPF-CHILLISMHHVOPRLH-----YKADENKGFCSKTAFTCTHPPETAFI-----AVT 215  
 Db 89 DEFTSPLTLNIPHCIPYIEKDKLKEIVNONNOSCFPMATIMPYIHLKPSILALNCMAVY 148  
 QY 216 STONKH-----ITQLTENNPFAKFGSDMLHNRSM-----Q 251  
 Db 149 ITRINKUIVSVTFHFGRELTLELDALEBEAF-----DEFTHGISOFEPYDHLLGYMKA 204  
 QY 252 SNEYVY-----PRSYRQKVAHNSPSSPSR-----ATLSNLSNGO 263  
 Db 205 SLERPEVILFLKYEYDKKDKPTSNVRRLAEPIGYFFFEEREKGVIESIKLCSFENLSM- 293  
 QY 291 YOCENGSSGSDQLLPPN 309  
 Db 264 --LEVNNSSNGSKGLPTEN 280

```

US: RESULT 13
US-09-908-193-19
Sequence 19, Application US/09908193
Publication No. US20020192746A1
INVENTOR: RASTELLI, LUCA
APPLICANT: RASTELLI, LUCA
APPLICANT: SHIMKETS, RICHARD A.
APPLICANT: ZERHUSEN, BRYAN
APPLICANT: MALYANKAR, BRIEL M.
APPLICANT: PADIGAN, MORALIDHARA
TITLE OF INVENTION: NOVEL POLYMUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
CURRENT FILING DATE: 2001-07-18
CURRENT APPLICATION NUMBER: US-09/908,193
PRIOR FILING DATE: 2000-07-24
PRIOR APPLICATION NUMBER: 60/220,273
PRIOR FILING DATE: 2000-07-24
PRIOR APPLICATION NUMBER: 60/221,650
PRIOR FILING DATE: 2000-07-28/221,233
PRIOR FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/220,912
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/218,875
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/218,870
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/218,901
PRIOR FILING DATE: 2000-07-18
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 1252
TYPE: CDS
ORGANISM: Mus musculus
US-09-908-193-19

Query Match 4.9% Score 92.5; Db 9; Length 1252;
Best Local Similarity 20.7%; Pred. No. 14;
Matches 92; Conservative 44; Mismatches 126; Indels 183; Gaps 22;
11 ANTPEPDKATKDCSKSEKALGASKSSSSPDAAATQOQICATYHEKELMKPEY 70
Db 626 SHVPAPD-DELKYRAKMEELVSNOPPHITQ-----ISQYKL-----MREY 667
OY 71 GTEMI---TANGRRFYSKYKATYGLNKRTKTLANDIYPADHRKRF-ADNK----- 120
Db 668 GEEADSDGPPRGGRDDANDGDGPRLEKRYKQYETLQVPGRLYEKVLAKNNHEDGYA 127
OY 121 -HWYCKRKE-----PAMGRILYVHPDPATGACNNR-OLVSFGKATKLTN----- 162
Db 728 AAWK--GKTEAEPDPLPIQGRPLP-PAHVHAESNSTSILWKKKPEFTTVYIVANYTV 784
OY 163 -----WN-----LD-PGCHILNMSMKY 179
Db 785 REGPAGRNALSVYTTTSSGDLIGLGEKPTKEFVAOSHGVDDPDQGVSSVNSLTPD 844

```



Tue Jun 17 10:11:34 2003

us-10-027-059a-1.rapb

Page 8

Db	463	LDSPDPSRPNSSNLSITKGRS	-----SGSS-----	PTRRTRVCVCHSS	505
QY	327	KGECDP-WS	335		
Db	506	DSPCMRPSMS	515		

Search completed: June 17, 2003, 10:03:51  
Job time : 22 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 09:44:00 ; Search time 19 Seconds  
(without alignments)  
1765.838 Million cell updates/sec

Title: US-10-027-059A-1

Perfect score: 1889

Sequence: 1 MAADRGFLAHTPEPPAK.....CDHFWSCFLSLFLSLGNG 349

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PTR.73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	958	50.7	184	2 S72231	transcription fact
2	865.5	45.8	173	2 S72230	transcription fact
3	658.5	34.9	423	2 B88450	protein F21H11.3 (
4	649.5	34.4	701	1 S46458	transcription fact
5	633.5	33.5	702	2 G01840	T-box protein 2 -
6	629.5	33.3	540	2 S72233	transcription fact
7	620	32.8	974	1 A0213	optic lobe develop
8	581	30.8	182	2 D56530	Ce-tbx-2 protein -
9	550.5	29.1	252	2 A88508	protein H14A12.4 (
10	549	29.1	344	2 A88023	protein T27A1.6 (l
11	544.5	28.8	436	1 S08156	gene T protein - m
12	541.5	28.7	432	1 A41056	brachyury homolog
13	536.5	28.4	423	1 A49135	gastrulation regul
14	536.5	28.4	681	2 I78558	hypothetical Brach
15	486	25.7	720	1 A55160	Trg protein - frui
16	474.5	25.1	664	2 T29011	hypothetical prote
17	419	22.2	182	2 C56530	Ce-tbx-7 protein -
18	391	20.7	479	1 S41015	transcription fact
19	384.5	20.4	342	2 C88579	protein tbx-8 (imp
20	381.5	20.2	377	1 S41019	transcription fact
21	340	18.0	352	2 T31544	hypothetical prote
22	282.5	15.0	303	2 T19424	hypothetical prote
23	267.5	14.2	322	2 E8473	protein F40H6.4 (l
24	221.5	11.7	376	2 T15166	hypothetical prote
25	171.5	9.1	302	2 T31542	hypothetical prote
26	141.5	7.5	507	2 T26809	hypothetical prote
27	133.5	7.1	331	2 T28023	hypothetical prote
28	123	6.5	414	2 T33057	hypothetical prote
29	123	6.5	414	2 E89467	protein ZK380.1 (l

hypothetical prote  
conserved hypothet  
hypothetical prote  
hypothetical prote  
probable retroviri  
transmembrane rece  
hypothetical coile  
T50401  
fork head protein  
Bifunctional phosph  
probable retroelem  
Bifunctional phosph  
probable membrane  
glial growth facto  
regulatory protein  
protein kinase C (

#### ALIGNMENTS

##### RESULT 1

S72231  
Transcription factor tbx5 - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 04-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Nov-1999  
C:Accession: S72231  
R:Agulnik, S.; Garvey, N.; Hancock, S.; Ruvinsky, I.; Chapman, D.L.; Agulnik, I.; B  
Genetics 144, 249-254, 1996  
A:Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.  
A:Reference number: S72230; MUID: 97032942; PMID: 8878690  
A:Accession: S72231  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-184 <AGU>  
A:Cross-references: SMI:U57230; NID: g1620599; PIDN: AAC53109.1; PID: g1620600  
A:Note: residue 184 is not shown  
C:Genetics:  
C:A gene: tbx5  
C:Superfamily: T-box homology  
C:Keywords: DNA binding  
F:1-184/Domain: T-box homology (fragment) <TBX>

Query Match 50.7%; Score 958; DB 2; Length 184;  
Best Local Similarity .967%; Pred. No. 1.9e-70;  
Matches 178; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 63 LWLKFEVGTGEMITTKAGRRMFPSPVKVYVGLNPKTKYLLMDLVPADDRYKFDANKWS 122  
|||||  
Db 1 LWLKFRSGVTEMITTKAGRRMFPSPVKVYVGLNPKTKYLLMDLVPADDRYKFDANKWS 60  
QY 123 VTGKAEPAMPGRLYVHPDSPAFAHMRQLVSVFQKLKLTNNHLDVFGHILNHNHXYQPR 182  
|||||  
Db 61 VTGKAEPAMPGRLYVHPDSPAFAHMRQLVSVFQKLKLTNNHLDVFGHILNHNHXYQPR 120  
QY 183 LHIVKADENNGFGSNTAFCTHVPETAFIVTSYQNHKITQLKIENNPFAKFGSGSDM 242  
|||||  
Db 121 LHIVKADENNGFGSNTAFCTHVPETAFIVTSYQNHKITQLKIENNPFAKFGSGSD 180  
QY 243 ELHR 246  
|||||  
Db 181 ELHR 184

##### RESULT 2

S72230  
Transcription factor tbx4 - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 04-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Nov-1999  
C:Accession: S72230  
R:Agulnik, S.I.; Garvey, N.; Hancock, S.; Ruvinsky, I.; Chapman, D.L.; Agulnik, I.; B  
Genetics 144, 249-254, 1996  
A:Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.

A:Reference number: 872230; MUID:97032942; PMID:8878690

A:Accession: S72230

A:Status: nucleic acid sequence not shown

A:Residues: 1-173 <AGU>

A:Cross-references: EMBL:U57329; NID:q1620597; PIDN:AA53108.1; PID:q1620598

A:Note: the sequence of residues 172-173 and the corresponding nucleic acid sequence are

C:Genetics:

A:Gene: tbx4

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:1-173/Domain: T-box Homology (fragment) <TBX>

Query Match 45.8%; Score 865.5; DB 2; Length 173;

Best Local Similarity 93.6%; Pred. No. 5.6e-63;

Matches 160; Conservative 5; Mismatches 1; Gaps 1;

74 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 133

1 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 59

134 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNG 193

60 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNA 119

194 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 244

120 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 170

RESULT 3

protein F21h1.3 [Imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

A:Accession: B88450

A:Residues: 1-423 <STO>

A:Cross-references: GB:chr\_III; PIDN:AA19434.1; PID:q506796; GSPDB:GN00021; CESP:F21h1.3

A:Map position: 3

Query Match 34.9%; Score 658.5; DB 2; Length 423;

Best Local Similarity 45.3%; Pred. No. 1.1e-45;

Matches 144; Conservative 47; Mismatches 84; Indels 43; Gaps 9;

30 SALGAPSKSPSPQAAFT---QOCM-----EGI---KVFLEHRELWLKFEVGT 72

20 AGVGCGAGGPPNLFPSMLQAGFPFGVSPEDDGVTDPKVLEDELRELWQOFSGCT 79

73 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 132

80 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 139

133 GRLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADEN 191

140 KTLVTHPSDPSTGEHMSKGNFHKLTNNISDKHYTILNSMHKYOPRLHYKADEN 199

192 NGFSKMTAFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 245

200 NLWVS---TFTVTFVTFETAFVTSYONHKITOLKIENNPFKRGSDMDL 256

246 RNSRMSQSKYVPVPRSTVRKVAS--NHSFSSFSRSLSTSSNL-----GSQYQCEGY 297

A:Reference number: 872230; MUID:97032942; PMID:8878690

A:Accession: S72230

A:Status: nucleic acid sequence not shown

A:Residues: 1-173 <AGU>

A:Cross-references: EMBL:U57329; NID:q1620597; PIDN:AA53108.1; PID:q1620598

A:Note: the sequence of residues 172-173 and the corresponding nucleic acid sequence are

C:Genetics:

A:Gene: tbx4

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:1-173/Domain: T-box Homology (fragment) <TBX>

Query Match 45.8%; Score 865.5; DB 2; Length 173;

Best Local Similarity 93.6%; Pred. No. 5.6e-63;

Matches 160; Conservative 5; Mismatches 1; Gaps 1;

74 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 133

1 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 59

134 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNG 193

60 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNA 119

194 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 244

120 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 170

RESULT 3

protein F21h1.3 [Imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

A:Accession: B88450

A:Residues: 1-423 <STO>

A:Cross-references: GB:chr\_III; PIDN:AA19434.1; PID:q506796; GSPDB:GN00021; CESP:F21h1.3

A:Map position: 3

Query Match 34.9%; Score 658.5; DB 2; Length 423;

Best Local Similarity 45.3%; Pred. No. 1.1e-45;

Matches 144; Conservative 47; Mismatches 84; Indels 43; Gaps 9;

30 SALGAPSKSPSPQAAFT---QOCM-----EGI---KVFLEHRELWLKFEVGT 72

20 AGVGCGAGGPPNLFPSMLQAGFPFGVSPEDDGVTDPKVLEDELRELWQOFSGCT 79

73 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 132

80 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 139

133 GRLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADEN 191

140 KTLVTHPSDPSTGEHMSKGNFHKLTNNISDKHYTILNSMHKYOPRLHYKADEN 199

192 NGFSKMTAFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 245

200 NLWVS---TFTVTFVTFETAFVTSYONHKITOLKIENNPFKRGSDMDL 256

246 RNSRMSQSKYVPVPRSTVRKVAS--NHSFSSFSRSLSTSSNL-----GSQYQCEGY 297

A:Reference number: 872230; MUID:97032942; PMID:8878690

A:Accession: S72230

A:Status: nucleic acid sequence not shown

A:Residues: 1-173 <AGU>

A:Cross-references: EMBL:U57329; NID:q1620597; PIDN:AA53108.1; PID:q1620598

A:Note: the sequence of residues 172-173 and the corresponding nucleic acid sequence are

C:Genetics:

A:Gene: tbx4

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:1-173/Domain: T-box Homology (fragment) <TBX>

Query Match 45.8%; Score 865.5; DB 2; Length 173;

Best Local Similarity 93.6%; Pred. No. 5.6e-63;

Matches 160; Conservative 5; Mismatches 1; Gaps 1;

74 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 133

1 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 59

134 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNG 193

60 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNA 119

194 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 244

120 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 170

RESULT 3

protein F21h1.3 [Imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

A:Accession: B88450

A:Residues: 1-423 <STO>

A:Cross-references: GB:chr\_III; PIDN:AA19434.1; PID:q506796; GSPDB:GN00021; CESP:F21h1.3

A:Map position: 3

Query Match 34.9%; Score 658.5; DB 2; Length 423;

Best Local Similarity 45.3%; Pred. No. 1.1e-45;

Matches 144; Conservative 47; Mismatches 84; Indels 43; Gaps 9;

30 SALGAPSKSPSPQAAFT---QOCM-----EGI---KVFLEHRELWLKFEVGT 72

20 AGVGCGAGGPPNLFPSMLQAGFPFGVSPEDDGVTDPKVLEDELRELWQOFSGCT 79

73 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 132

80 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 139

133 GRLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADEN 191

140 KTLVTHPSDPSTGEHMSKGNFHKLTNNISDKHYTILNSMHKYOPRLHYKADEN 199

192 NGFSKMTAFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 245

200 NLWVS---TFTVTFVTFETAFVTSYONHKITOLKIENNPFKRGSDMDL 256

246 RNSRMSQSKYVPVPRSTVRKVAS--NHSFSSFSRSLSTSSNL-----GSQYQCEGY 297

A:Reference number: 872230; MUID:97032942; PMID:8878690

A:Accession: S72230

A:Status: nucleic acid sequence not shown

A:Residues: 1-173 <AGU>

A:Cross-references: EMBL:U57329; NID:q1620597; PIDN:AA53108.1; PID:q1620598

A:Note: the sequence of residues 172-173 and the corresponding nucleic acid sequence are

C:Genetics:

A:Gene: tbx4

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:1-173/Domain: T-box Homology (fragment) <TBX>

Query Match 45.8%; Score 865.5; DB 2; Length 173;

Best Local Similarity 93.6%; Pred. No. 5.6e-63;

Matches 160; Conservative 5; Mismatches 1; Gaps 1;

74 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 133

1 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 59

134 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNG 193

60 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNA 119

194 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 244

120 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 170

RESULT 3

protein F21h1.3 [Imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

A:Accession: B88450

A:Residues: 1-423 <STO>

A:Cross-references: GB:chr\_III; PIDN:AA19434.1; PID:q506796; GSPDB:GN00021; CESP:F21h1.3

A:Map position: 3

Query Match 34.9%; Score 658.5; DB 2; Length 423;

Best Local Similarity 45.3%; Pred. No. 1.1e-45;

Matches 144; Conservative 47; Mismatches 84; Indels 43; Gaps 9;

30 SALGAPSKSPSPQAAFT---QOCM-----EGI---KVFLEHRELWLKFEVGT 72

20 AGVGCGAGGPPNLFPSMLQAGFPFGVSPEDDGVTDPKVLEDELRELWQOFSGCT 79

73 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 132

80 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 139

133 GRLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADEN 191

140 KTLVTHPSDPSTGEHMSKGNFHKLTNNISDKHYTILNSMHKYOPRLHYKADEN 199

192 NGFSKMTAFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 245

200 NLWVS---TFTVTFVTFETAFVTSYONHKITOLKIENNPFKRGSDMDL 256

246 RNSRMSQSKYVPVPRSTVRKVAS--NHSFSSFSRSLSTSSNL-----GSQYQCEGY 297

A:Reference number: 872230; MUID:97032942; PMID:8878690

A:Accession: S72230

A:Status: nucleic acid sequence not shown

A:Residues: 1-173 <AGU>

A:Cross-references: EMBL:U57329; NID:q1620597; PIDN:AA53108.1; PID:q1620598

A:Note: the sequence of residues 172-173 and the corresponding nucleic acid sequence are

C:Genetics:

A:Gene: tbx4

C:Superfamily: T-box homology

C:Keywords: DNA binding

F:1-173/Domain: T-box Homology (fragment) <TBX>

Query Match 45.8%; Score 865.5; DB 2; Length 173;

Best Local Similarity 93.6%; Pred. No. 5.6e-63;

Matches 160; Conservative 5; Mismatches 1; Gaps 1;

74 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 133

1 MITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 59

134 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNG 193

60 RLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADENNA 119

194 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 244

120 FGSKNATFCTHVFFETAFVTSYONHKITOLKIENNPFKRGSDMDL 170

RESULT 3

protein F21h1.3 [Imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

A:Accession: B88450

A:Residues: 1-423 <STO>

A:Cross-references: GB:chr\_III; PIDN:AA19434.1; PID:q506796; GSPDB:GN00021; CESP:F21h1.3

A:Map position: 3

Query Match 34.9%; Score 658.5; DB 2; Length 423;

Best Local Similarity 45.3%; Pred. No. 1.1e-45;

Matches 144; Conservative 47; Mismatches 84; Indels 43; Gaps 9;

30 SALGAPSKSPSPQAAFT---QOCM-----EGI---KVFLEHRELWLKFEVGT 72

20 AGVGCGAGGPPNLFPSMLQAGFPFGVSPEDDGVTDPKVLEDELRELWQOFSGCT 79

73 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 132

80 EMITKAGRMFPYSKYVKTGLNPKTKYLLMDIVPADDRHYKFAFNKMSYTGKAEFAMFG 139

133 GRLYVHPSDPATCAHWRQVLSFOKLKLTNNHLDPEGHILNSMHKYOPRLHYKADEN 191

A:Map position: 17q21-17q22  
 C:Superfamily: mouse transcription factor tbx2; T-box homology  
 F:104-285/Domain: T-box homology <TBX>

Query Match 33.5%; Score 633.5; DB 2; Length 702;  
 Best Local Similarity 52.0%; Pred. No. 2.4e-43;  
 Matches 130; Conservative 30; Mismatches 53; Indels 37; Gaps 3;

QY 23 PCDSKPSALGAPKSP-----SSPOAFTQCG----- 51  
 DB 29 PALAUPGALAKLPDPGLAGAAAAAANGLVHVSALGPHPPAAHLRLSKLEPE 88  
 QY 52 -----EGIKVFLHERELMLKFEVGTMTITKAGRMFSPYKVKVGLNPKTKYILLMDIV 107  
 DB 89 DEVEDPKVTLKLELMDWDFHLGTETMTKSGRMFPFPRKVRVSGLDKRAKYILLMDIV 148  
 QY 108 PADHURKADNKSVTGAEPAMPGLRYVHPDSPATGAHMRQLVSPFKLKTNNHLDLP 167  
 DB 149 ADDCRKATFNSWVAGKADPEPRMTYIHPSDPATGEGMAKPVAFKALKUTNNSDK 208  
 QY 168 FGHILNSMKHYQPRHLIVKADENKFGSKTAFCTHVPFETAFVTSQNHKTKOLKI 227  
 DB 209 HGFTILNSMKHYQPRHLIVRA--NDILKLPYFRTYVFPETOFIVATYQNKTIOLKI 266  
 QY 228 ENNPFAKGR 237  
 DB 267 DNNPFAKGR 276

RESULT 6  
 572233  
 transcription factor tbx6 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 05-Nov-1999  
 C:Accession: S72233  
 R:Aquilnik, S.I.; Chapman, D.L.; Hancock, S.; Silver, L.M.  
 submitted to the EMBL Data Library, May 1996  
 A:Reference number: 572233  
 A:Accession: S72233  
 A:Molecule type: mRNA  
 A:Residues: 1-540 <AGU>  
 A:Cross-references: EMBL:U57331; NID:q1620601; PIDN:AC53110.1; PID:g1620602  
 R:Aquilnik, S.I.; Garvey, N.; Hancock, S.; Ruvinsky, I.; Chapman, D.L.; Aguilnik, I.; Bohl  
 Genetics 144, 249-254, 1996  
 A:Title: Evolution of mouse T-box genes by tandem duplication and cluster dispersion.  
 A:Reference number: 572230; MUID:97032942; PMID:8878690  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 100-280 <AGX>  
 A:Cross-references: EMBL:U57331  
 C:Genetics:  
 A:Gene: tbx6  
 C:Superfamily: T-box homology  
 C:Keywords: DNA binding  
 F:100-282/Domain: T-box homology <TBX>

Query Match 33.3%; Score 629.5; DB 2; Length 540;  
 Best Local Similarity 43.4%; Pred. No. 3.6e-43;  
 Matches 138; Conservative 45; Mismatches 80; Indels 55; Gaps 9;

QY 11 AHTLEPAKDLPCDSKPSALGAPKSPSPQAFTQCGMEGIVFLHERELMLKFEV 70  
 DB 62 AAAPLPSPS-----PIALG-PETAPPPPAL---HSLPGVSLSEQLWKEFSAV 107  
 QY 71 GTEMLITKAGRMFSPYKVKVGLNPKTKYILLMDIVPADHRYKADNKWSVTGAEP 130  
 DB 108 GTEMLITKAGRMFSPYKVKVGLNPKTKYILLMDIVPADHRYKADNKWSVTGAEP 167  
 QY 131 MPGRLYVHPDSPATGAHMRQLVSPFKLKTNNHLDLPFGHILNSMKHYQPRHLIVRADE 190  
 DB 168 LPDRYTHIPDSPATGAHMRQVSPFHRVKLTNSTLDPCHLILHSMHYQPRHLIVRA-- 225

QY 191 NNGFGSKNTAFCTH-----VFPETAFIVTSYQNHKITOLKIENNPAKGR-GSDD 241  
 DB 226 -----TQLCSQHGCVASFPREPPTTISVTAYONPRITOLKIANNPAKGRNGRN 277  
 QY 242 MELHMRMQSKSEYVYVPRSTVRQVASHSPFSSSRALSTSSNLGSOQVCNCGVSGPS 301  
 DB 278 CRREDARVKKRLAGPEPVAN--EACGSDTPGSG-----PCDSTLGDI 319  
 QY 302 QDLLPPNPYP-----LPQ 315  
 DB 320 RSDDEQAPTCKLLLPQ 337

RESULT 7  
 A40213  
 optic lobe development omb protein - fruit fly (Drosophila melanogaster)  
 N:Alternate names: omb protein  
 C:Species: Drosophila melanogaster  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A40213; S34827  
 R:Flugfelder, G.O.; Roth, H.; Poock, B.; Kerscher, S.; Schwarz, H.; Jonschker, B.; H  
 Proc. Natl. Acad. Sci. U.S.A. 89, 1199-1203, 1992  
 A:Title: The *techa*(1) optomotor-blind gene of *Drosophila melanogaster* is a major orga  
 A:Reference number: A40213; MUID:92159016; PMID:1741374  
 A:Accession: A40213  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-974 <CFL>  
 A:Cross-references: GB:M81796; NID:g158018; PIDN:AAA28736.1; PID:g158019  
 A:Note: sequence extracted from NCBI backbone (NCBI:82036)  
 R:Poock, B.; Balles, J.; Flugfelder, G.O.  
 Mol. Gen. Transc. 238, 325-332, 1993  
 A:Title: Transcriptional regulation in the optomotor-blind locus of *Drosophila melanog*  
 A:Reference number: A34827; MUID:93261414; PMID:8492800  
 A:Accession: S34827  
 A:Molecule type: DNA  
 A:Residues: 1-947 <ROE>  
 A:Cross-references: GB:S61732; NID:g402317; PIDN:AA826697.1; PID:g402318  
 A:Experimental source: larva  
 C:Genetics:  
 A:Gene: FlyBase:bi  
 A:Cross-references: FlyBase:Fgn000179  
 C:Superfamily: optic lobe development omb protein; T-box homology  
 C:Keywords: DNA binding  
 F:337-521/Domain: T-box homology <TBX>

Query Match 32.8%; Score 620; DB 1; Length 974;  
 Best Local Similarity 46.7%; Pred. No. 4.7e-42;  
 Matches 135; Conservative 38; Mismatches 70; Indels 46; Gaps 6;

QY 55 KVFLEHERELMLKFEVGTMTITKAGRMFSPYKVKVGLNPKTKYILLMDIVPADHRY 114  
 DB 329 KVTLEGLDLWEKFLKLTETMTKSGRMFPFPRKVRVSGLDKRAKYILLMDIVPADHRY 388  
 QY 115 KEADNKSVTGAEPAMPGLRYVHPDSPATGAHMRQLVSPFKLKTNNHLDLPFGHI 171  
 DB 389 KEHNSRMVAGKADPEPRMTYIHPSDPATGEGMAKPVAFKALKUTNNSDKHGFVSTT 448  
 QY 172 ILNSMKHYQPRHLIVKADENKFGSKNTAFCTHVPFETAFVTSQNHKTKOLKIENN 231  
 DB 449 ILNSMKHYQPRHLIVRA--NDILKLPYFRTYVFPETOFIVATYQNKTIOLKIENN 506  
 QY 232 FAKGRF--GSDDMELHMRMQSKSEYVYVPRSTVRQVASHSPFSSSRALSTSSNLGS 289  
 DB 507 FAKGLRDTGAGKRE-----RNCVQALMSNR---GSOSKLNPHVSSS 547  
 QY 290 QVCNCGVSGPSQDLLP-----PPNPFLPQEHS 318  
 DB 548 RAPLHLHGAGRPHLPHALLDNODDOKLLDVGVGPPQSPLLPLSHS 596

RESULT 8  
 D56530

ce-tbx-2 protein - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans  
C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 24-Jul-1997  
C:Accession: D56530  
R:Agulnik, S.I.; Bollag, R.J.; Silver, L.M.  
A:Title: Conservation of the T-box gene family from Mus musculus to Caenorhabditis elegans  
A:Accession: D56530 #sequence\_revision 19-Oct-1995 #text\_change 24-Jul-1997  
A:Accession: D56530 #sequence\_revision 19-Oct-1995 #text\_change 24-Jul-1997  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-182 <AGU>  
A:Cross-references: GB:U11279  
C:Superfamily: T-box homology  
C:Keywords: DNA binding  
F:1-182/Domain: T-box homology <TBX>

Query Match 30.8%; Score 581; DB 2; Length 182;

Best Local Similarity 60.8%; Pred. No. 7.1e-40;  
Matches 107; Conservative 31; Mismatches 34; Indels 4; Gaps 2;

QY 63 LMLKHEVGTMTAKGRMPSPYKVKYGLNPKTKYLLMDIVPADDHYKFAADKNS 122  
Db 1 LMQQFSGCTENYKSGRIIPAYRVKISGLDKSQYFVNDLVADPEHYKFNKSRW 60  
QY 123 VTGKAEPAMPGRLYVHIDPSATGAHMRQLVSFOKLITNNHLDPFGHILNSMKHYOPR 182  
Db 61 IAGKADPEPKTYLIHPDSTGEGHMSKGANFKLITNNISDKHGTYILNSMKHYOPR 120  
QY 183 LHIYK-ADENNGRGSKNAPCTCHVPTAFATNTSYONHKITLKNPFAKGR 237  
Db 121 LHVRCADRLNLYS---TPTFVFRETETATVAYONKAVTELKLNPFANPFAKGR 173

#### RESULT 9

protein H14A12.4 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-May-2001  
C:Accession: A88023  
R:Anonymous. The C. elegans Sequencing Consortium.  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see webistes genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 282, 2012-2018, 1998  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-252 <STO>  
A:Cross-references: GB:chr.IIIT; PID:NAB70985.1; PID:g2429447; GSPDB:GN00021; CESP:H14A12.4  
A:Note: strong similarity to Brachyury-like proteins  
C:Genetics:  
A:Gene: H14A12.4  
A:Map position: 3

Query Match 29.18%; Score 550.5; DB 2; Length 252;  
Best Local Similarity 53.9%; Pred. No. 3.3e-37;  
Matches 104; Conservative 31; Mismatches 49; Indels 9; Gaps 3;  
QY 47 TQGMGKIKVFLHRELMLKPFHEVGTMTAKGRMPSPYKVKYGLNPKTKYLLMDI 106  
Db 24 TRNCKRFLRVLOSSULMRFNIGLTHIVYKSGRMPEPTLSIAGDPVKSVMVVDL 83  
QY 107 --VPADDPHYKFAADKNSVTGKAEPAMPGRLYVHIDPSATGAHMRQLVSFOKLITNNH 164  
Db 84 ETEKMKRFYSFPHOSKWLSTGPESELPGRMVFHSDSPARAHMRAPSPDKMKLTNNQ 143  
QY 165 LDPEGHILNSMKHYOPRHLHVADENNGRGSKNAPCTCHVPTAFATNTSYONHKIT 224  
Db 144 LNNCHILNSMKHYOPRHLHVADENNGRGSKNAPCTCHVPTAFATNTSYONHKIT 224  
QY 225 LKTNPNFAKGR 237

Db 197 LKIESNPFAGFR 209

#### RESULT 10

protein T27A1.6 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: A88023  
R:Anonymous. The C. elegans Sequencing Consortium.  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see webistes genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 282, 2012-2018, 1998  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-344 <STO>  
A:Cross-references: GB:chr.II; PID:AB71046.1; PID:g2429517; GSPDB:GN00020; CESP:T27A1.6  
A:Note: strong similarity to brachyury-like proteins (T protein)  
C:Genetics:  
A:Gene: T27A1.6  
A:Map position: 2  
C:Superfamily: Caenorhabditis elegans transcription factor tbx9; T-box homology

Query Match 29.18%; Score 549; DB 2; Length 344;

Best Local Similarity 45.8%; Pred. No. 6.6e-17;  
Matches 115; Conservative 33; Mismatches 87; Indels 16; Gaps 5;

QY 18 DAKDLPCDSKPESALGADSKSPSPQAAF-TQGMGKIKVFLHRELMLKPFHEVGTMT 76  
Db 43 DVDVDLSSIPSr---SPERSGRPKTGURMKGNLPIECKLEGSELMKAFDGLGTEMI 99  
QY 77 TTAGRMFPSPYKVKYGLNPKTKYLLMDIVPADDHYKFAADK--HSVTGKAEPAMPGR 134  
Db 100 TSSGRMFVTVKSTNVLDAIYIFLDVVPVDSKRTIYVKSAMLTACKAEPVKNR 159  
QY 135 LVVHPDSPATGAHMRQLVSFOKLITNNHLDPFGHILNSMKHYOPRHLHVADENNGF 194  
Db 160 YLHPDSPFGDGLKLVHISFKETKLTNNVDKTHGLNSMKHYOPRHLHVADENNGF 219  
QY 195 GSKNTAF-----CTHVPTAFATNTSYONHKITLKNPFAKGR-----RGSDDMEL 244  
Db 220 DKNVNSEKHCTYTFETOPMAVAYONLITLKLTKENPNKPFGRDPTGRSPOEMER 279  
QY 245 HNSRMQSKKEY 255  
Db 280 SPQDMHLSNFY 290

#### RESULT 11

gene T protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S08156  
R:Herrmann, B.G.; Labelle, S.; Pousetka, A.; King, T.R.; Lehrach, H.  
A:Title: Cloning and characterization of the T protein, a transcription factor required in the mouse.  
A:Reference number: S08156; MUID:90158787; PMID:2154694  
A:Accession: S08156  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-436 <HER>  
A:Cross-references: GB:X51683; MUID:g55053; PIDN:CAA35985.1; PID:g55054  
C:Genetics:  
A:Map position: 17  
C:Superfamily: mouse gene T protein; T-box homology  
C:Keywords: DNA binding; transcription regulation  
F:51-227/Domain: T-box homology <TBX>

Query Match 28.8%; Score 544.5; DB 1; Length 436;  
 Best Local Similarity 43.8%; Pred. No. 2.1e-36;  
 Matches 117; Conservative 39; Mismatches 88; Indels 23; Gaps 5;

QY 54 IKVFLHERELMWLFKEVGTETMITKAGRMFESYKVKYTLGNPKTKYILLMDIVPADDRH 113  
 DB 42 LRVGLSEELMWLFKEVGTETMITKAGRMFESYKVKYTLGNPKTKYILLMDIVPADDRH 101  
 QY 114 YKPADNKWSTGKAEPMGRGLYVHPDSPATGAGHMRQLVDFQKALTNHLDPFQGHIL 173  
 DB 102 WKYVNGEWPVGKPPQAPSCVYIHPDSPNFGAHMKAPVFSKVLTKN-MNGGQQLM 160  
 QY 174 NSHHKYQPLRHIVKADENNGFSGKNTAFCTHVFFETAFIANTVTSQNHKITQKIENPFA 233  
 DB 161 NSHHKYQPLRHIVKADENNGFSGKNTAFCTHVFFETAFIANTVTSQNHKITQKIENPFA 214  
 QY 234 KGRFGSDMLHRM-MNSHOSKEYP-----GGTQRMITSHTCFPTQFIANTVTSQNHKITQKIENPFA 284  
 DB 215 KAFDADKRNHDKVMEPCQKQYSGQWGLYFGNGLTCLFPASSHTQFGG-----S 267  
 QY 285 NLGSOYOCENGSGPSODLLPPNY 311  
 DB 268 LSLPSTHCCERYPALNRHSSPYSPY 294

RESULT 12  
 A41056  
 Brachyury homolog - African clawed frog  
 N:Alternate names: gene T homolog  
 C:Species: Xenopus laevis (African clawed frog)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A41056  
 R:Smith, J.C.; Price, B.M.J.; Green, J.B.A.; Welgel, D.; Herrmann, B.G.  
 Cell 67, 79-87, 1991  
 A>Title: Expression of a Xenopus homolog of Brachyury (T) is an immediate-early response  
 A:Reference number: A41056; MUID:92005698; PMID:1717160  
 A:Accession: A41056  
 A:Molecule type: mRNA  
 A:Residues: 1-432 <SM>  
 A:CROSS-references: GB:M77243; NID:g213999; PIDN:AAA49663.1; PID:g214000  
 A:Comment: Both basic fibroblast growth factor (bFGF) and activin A induce expression of  
 A:Gene: Xbra  
 C:Superfamily: mouse gene T protein; T-box homolog  
 C:Keywords: DNA binding; mesoderm; transcription regulation  
 F:49-225/Domain: T-box homolog <TBX>

Query Match 28.7%; Score 541.5; DB 1; Length 432;  
 Best Local Similarity 41.2%; Pred. No. 3.6e-36;  
 Matches 113; Conservative 43; Mismatches 81; Indels 37; Gaps 5;

QY 54 IKVFLHERELMWLFKEVGTETMITKAGRMFESYKVKYTLGNPKTKYILLMDIVPADDRH 113  
 DB 40 LRVGLSEELMWLFKEVGTETMITKAGRMFESYKVKYTLGNPKTKYILLMDIVPADDRH 99  
 QY 114 YKPADNKWSTGKAEPMGRGLYVHPDSPATGAGHMRQLVDFQKALTNHLDPFQGHIL 173  
 DB 100 WKYVNGEWPVGKPPQAPSCVYIHPDSPNFGAHMKAPVFSKVLTKN-MNGGQQLM 158  
 QY 174 NSHHKYQPLRHIVKADENNGFSGKNTAFCTHVFFETAFIANTVTSQNHKITQKIENPFA 233  
 DB 159 NSHHKYQPLRHIVKADENNGFSGKNTAFCTHVFFETAFIANTVTSQNHKITQKIENPFA 212  
 QY 234 KGRFGSDMLHRM-MNSHOSKEYPVPVSTVROKVASNHSPSSSRALSTS-----285  
 DB 213 KAP-----LDKERNDY-----KDLIDGIDDSQHSNFSOLGTWLPNGGSLCSPN 258  
 QY 286 -----NLGSOYOCENGSGPSODLLPPNY 311  
 DB 259 HTQFGAPLSLSSPHGCCERYSSLRNHSAPSPY 292

RESULT 13  
 A49125  
 gastrulation regulatory protein zif-T - zebra fish  
 C:Species: Brachydanio rerio (zebra fish)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A49125  
 R:Schulte-Merker, S.; Ho, R.K.; Herrmann, B.G.; Nusslein-Volhard, C.  
 Development 116, 1021-1032, 1992  
 A>Title: The protein product of the zebrafish homologue of the mouse T gene is expressed  
 A:Reference number: A49125; MUID:93201989; PMID:1295726  
 A:Accession: A49125  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-423 <SCH>  
 A:CROSS-references: GB:S57147; NID:g299316; PIDN:AAB25829.1; PID:g299317  
 A:Experimental source: embryo  
 A>Note: sequence extracted from NCBI backbone (NCBIN:127982, NCBIPI:127985)  
 C:Superfamily: mouse gene T protein; T-box homolog  
 F:44-220/Domain: T-box homolog <TBX>

Query Match 28.4%; Score 536.5; DB 1; Length 423;  
 Best Local Similarity 41.5%; Pred. No. 9e-36;  
 Matches 125; Conservative 41; Mismatches 90; Indels 45; Gaps 10;

QY 40 SSPQAAFTQOGME-----GKVFLEHERELMWLFKEVGTETMITKAGRMFESYKVKY 92  
 DB 15 SAVSEF-QKSGKDGASERDKLSDAELMTKFKELTNEMIVTKGRHFPVLRSYV 73  
 QY 93 GLNPKTKYILLMDIVPADDRHFKADNKVSTGKAEPMGRGLYVHPDSPATGAGHMRQL 152  
 DB 74 GLDPMNMYSLVLDFFVAADNRKRYVNGEWPVGKPPQAPSCVYIHPDSPNFGAHMKAP 133  
 QY 153 VSKOKLTKNHLDPFGHILNSHMKYQPLRHIVKADENNGFSGKNTAFCTHVFFETAFI 212  
 DB 134 VSESKVLSNK-LNGGQQLMNSLHKEPRHIIKVK-----GGIOKMISOSFPETQFI 186  
 QY 213 AVTSYQNHKITQKIENPFGKSGDMLHRMSKQKEYPVVSTVROKVASNH 272  
 DB 187 AVTAYQNEITALKIKHNPFAK-----LDKERSDHEVP-----DHS 226  
 QY 273 PFSSSRALSTSSNGSOYOCENGSGPSODLLPPN-NPYLPDSHQSOLYHCTKRGKGD 331  
 DB 227 TDNQO----SGYSLGGWFLPSNGMGPSSS--PPQNGAPV---HSSGSYCYERYSSLRN 277  
 QY 332 H 332  
 DB 278 H 278

RESULT 14  
 I78558  
 hypothetical Brachyury (T)-related transcription factor - mouse  
 C:Species: Mus sp. (mouse)  
 C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jul-2000  
 C:Accession: I78558  
 R:Bulfone, A.; Smiga, S.M.; Shimanura, K.; Peterson, A.; Puelles, L.; Rubenstein, J.L.  
 Neuron 15, 63-78, 1995  
 A>Title: T-brain-1: a homolog of Brachyury whose expression defines molecularly distinct  
 A:Reference number: 158171; MUID:95344783; PMID:7619531  
 A:Accession: I78558  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-681 <RES>  
 A:CROSS-references: GB:S78858; NID:g122544; PIDN:AAA92011.1; PID:g122545  
 A:Gene: Tbr-1/Tes-56  
 C:Genetics:  
 F:213-401/Domain: T-box homolog <TBX>

Query Match 28.4%; Score 536.5; DB 2; Length 681;  
 Best Local Similarity 38.7%; Pred. No. 1.7e-35;  
 Matches 136; Conservative 53; Mismatches 105; Indels 57; Gaps 13;

us-10-027-059a-1.rpr

Tue Jun 17 10:11:35 2003

```

OY 10 LAHTPLEPDA-KDLPDCSPESALGAPSKSPSSPO-----ANF-----TQGMEGI 54
Db 146 MAHPVITNGAYNSLSSSPQ---GYTAGPYPQYGVSHYQCAFTQFSSTQGLVPG 202
OY 55 K-VFLHERELMLKHEVGTGEMITKAGRRMFPYSKYKVTGLNPKTKYILLMDIVPADDH 112
Db 203 KAQYVLCNRPWLKFRHOTEMITTKOGRMFPFISNIGLDPTAHYNIQFVIVLADPN 262
OY 113 RYKADNWSVTGKAEFAPMG-RLYVHPSDPATGAHMRPOLYSFOKLKLTNNHLOPEGH- 170
Db 263 HRFQGGKWPVCKKADTVTQGNRVVYHPPSDPTNGAHHMRQEISETGKLKLTNNKGSANNNG 322
OY 171 --ILNSMHKYPQLRHIVKADENNGFSKNTA-----FCHVFPETAFIAVTSYONHKITQ 224
Db 323 QAVYLSLHKYQRLHVVVEVED--GTEDTSGPQGVQTFPTFQFIATVQNTDITQ 379
OY 225 LKINENPAKGRFGS-----DNELHRMSKOSKEYPVWP-----RSTVRQKVASN 270
Db 380 LKIDHNPFAKGRDNYDTIVTGCMDRLTPSPNDSPSOIVPGARYAMAGSFLOQOFVSN 439
OY 271 HS-----PFSSSNALSTSSNLGSOYOCEN-QVSGPSODLLPPPN 309
Db 440 YAKARHPGAGAGPGGTDRSVPHNTGILLSPOQAEDPGAPSPQRMVFTPAN 490

RESULT 15
A:55160
A:Protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A55160
R:Kispart, A.; Herrmann, B.G.; Leptin, M.; Reuter, R.
Gene: Dec 8, 213 Dec 20, 1999
A:1. Homo. Of 230. mouse Brachyury gene are involved in the specification of poste
A:Reference number: A55160; MUID:95047361; PMID:7958884
A:Accession: A55160
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
C:Residues: 1-720 <KIS>
A:Gene: FlyBase:Trg
A:Cross-references: FlyBase:FBgn0011723
C:Superfamily: Drosophila Trg protein; T-box homology
F:96-272/Domain: T-box homology <TBX>

Query Match 25.7%; Score 486; DB 1: Length 720;
Best Local Similarity 35.2%; Pred. No. 2,3e-31;
Matches 102; Conservative 50; Mismatches 82; Indels 56; Gaps 6;

OY 54 IKVFLHERELMLKHEVGTGEMITTKAGRRMFPYSKYKVTGLNPKTKYILLMDIVPADDH 113
Db 87 LAISLDDREMLRFONLNEMIVTNGRRMFPVYKISASGLDPAAMYTLVLEFVQIDSHR 146
OY 114 YKFDANKSVYTKAEAPMGRLVYHPSDPATGAHMRPOLYSFOKLKLTNNHLDPPGHIL 173
Db 147 WKYVNGEVPFGKAEVPSNPYIVHPSDPNGAHHMRQEISETGKLKLTNNHLDPPGHIL 205
OY 174 NSMHKYPQLRHIVKADENNGFSKNTAFTCHVFPETAFIAVTSYONHKITQKINNPPA 233
Db 206 NSLHKYEPRLVLRV-----GSEORHVVTYFPFETQFIATVAYQNEEVTSLKIKYNPFA 259
OY 234 KPGSGSDMELHRMSKOSKEYP--VYPASTVROKVASNMHSPSSSESLSTSSNLGSOY 291
Db 260 KAF-----LDAKERDPTLYPHDTHYGLMPLPPHTYAAAAAFA----- 298
OY 292 QCENGVSQSDLLPPPNPYPQLPQHSQIYHC-----TKRKGEC 331
Db 299 -----PPPLSLAQSHGLVASCPSVSSAGSVGSPSGGSCD 332

Search completed: June 17, 2003, 09:49:27
Job time : 22 secs

```

LOCUS (LOC): HSU80987 GenBank (R)  
GenBank ACC. NO. (GBN): U80987  
GenBank VERSION (VER): U80987.1 GI:2281318  
CAS REGISTRY NO. (RN): 186848-85-1  
SEQUENCE LENGTH (SQL): 1050  
MOLECULE TYPE (CI): mRNA; linear  
DIVISION CODE (CI): Primates  
DATE (DATE): 26 Jul 1997  
DEFINITION (DEF): Human transcription factor TBX5 mRNA; complete cds.  
SOURCE: human.  
ORGANISM (ORGN): Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;  
Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini;  
Hominidae; Homo  
NUCLEIC ACID COUNT (NA): 278 a 301 c 248 g 223 t  
REFERENCE: 1 (bases 1 to 1050)  
AUTHOR (AU): Basson, C.T.; Bachinsky, D.R.; Lin, R.C.; Levi, T.;  
Elkins, J.A.; Soultz, J.; Grayzel, D.; Kroumpouzou, E.;  
Traill, T.A.; LeBlanc-Straceski, J.; Renault, B.;  
Kucherlapati, R.; Seidman, J.G.; Seidman, C.E.  
TITLE (TI): Mutations in human TBX5 [corrected] cause limb and  
cardiac malformation in Holt-Oram syndrome  
JOURNAL (SO): Nat. Genet., 15 (1), 30-35 (1997)

THIS PAGE BLANK (USPTO)